

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 13:45:48 ; Search time 2760 Seconds

(without alignments)  
5883.828 Million cell updates/sec

Title: US-09-720-451-5

Perfect score: 1 tgattgtccaccatattt.....tcagtcatgcygcacgtg 558

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_da:\*

2: gb\_hlg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_da:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_hlg\_hum:\*

31: em\_hlg\_inv:\*

32: em\_hlg\_other:\*

33: em\_hlg\_mus:\*

34: em\_hlg\_pln:\*

35: em\_hlg\_rtd:\*

36: em\_hlg\_mam:\*

37: em\_hlg\_vrt:\*

38: em\_sy:\*

39: em\_higo\_hum:\*

40: em\_higo\_mus:\*

41: em\_higo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	378.8	67.9	1897	8	AF181967	AF181967 Arabidops
2	378.8	67.9	1917	8	AY050434	AY050434 Arabidops
3	378.8	67.9	2034	8	AE370515	AE370515 Arabidops
4	377.2	67.6	2037	8	AT245414	AT245414 Arabidops
5	358.8	64.3	1810	8	AY122922	AY122922 Arabidops
6	358.8	64.3	1994	8	AF181966	AF181966 Arabidops
7	358.8	64.3	2002	8	AY070034	AY070034 Arabidops
8	358.8	64.2	2225	8	AF174486	AF174486 Zea mays
9	130	23.3	38571	3	U39849	U39849 Caenorhabdi
10	127.2	22.8	1785	12	AF441241	AF441241 Synthetic
11	126.8	22.7	81624	3	AC099763	AC099763 Caenorhab
12	121.8	21.8	109741	8	AC004005	AC004005 Arabidops
13	117.4	21.0	37093	8	SPAC56F8	SPAC56F8 S.pombe chr
14	116.4	20.9	1971	6	AX069359	AX069359 Sequence
15	116.4	20.9	2187	6	AR144956	AR144956 Sequence
16	116.4	20.9	2187	6	AX050441	AX050441 Sequence
17	116.4	20.9	2196	12	HSU09806	U09806 Synthetic c
18	116.4	20.9	2219	6	A47328	A47328 Sequence 3
19	116.4	20.9	2219	6	AR097969	AR097969 Sequence
20	116.4	20.9	2219	6	AR145811	AR145811 Sequence
21	116.4	20.9	2219	6	AX356174	AX356174 Sequence
22	116.4	20.9	2219	6	AX452351	AX452351 Sequence
23	116.4	20.9	2220	6	A47326	A47326 Sequence 1
24	116.4	20.9	2220	6	AR097968	AR097968 Sequence
25	116.4	20.9	2220	6	AR145810	AR145810 Sequence
26	116.4	20.9	2220	6	AX356172	AX356172 Sequence
27	116.4	20.9	2313	9	HSR237672	AJ237672 Homo sapi
28	112.6	20.2	100906	8	ATF24G16	AL138647 Arabidops
29	102.4	18.4	1250	10	RNF57049	U57049 Rattus norv
30	67	12.0	3082	8	SCYGL125W	Z72647 S.cerevisia
31	67	12.0	3082	8	SCSUP44GN	X94106 S.cerevisia
32	64.6	11.6	110000	2	LMFCCHR36_02	Continuation (3 of
33	56.2	10.1	15247	1	AE000740	AE000740 Aquifex a
34	56	10.0	551	4	AF239166	AF239166 Sus scrof
35	54	9.7	12141	1	AE014080	AE014080 Buchnera
36	53.4	9.6	347550	1	AP001118	AP001118 Buchnera
37	52.2	9.4	215050	1	AL646057	AL646057 Ralstonia
38	51.8	9.3	198728	10	MMETTER06	AF105993 Mus muscu
39	51.8	9.3	198728	10	AL606929	AL606929 Mouse DNA
40	50.4	9.0	10592	1	AE006058	AE006058 Pasteurel
41	50.2	9.0	181581	2	AC113917	AC113917 Rattus no
42	49.2	8.8	10428	1	AE006357	AE006357 Lactococc
43	49	8.8	12059	1	AE005886	AE005886 Caulobact
44	48.8	8.7	11166	1	U32823	U32823 Haemophilus
45	48.6	8.7	235	6	AX452358	AX452358 Sequence

## ALIGNMENTS

RESULT 1

AF181967

LOCUS

DEFINITION Arabidopsis thaliana methylentetrahydrofolate reductase MTHFR2

ACCESSION AF181967

VERSION AF181967.1

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 1897)

Rojes, S., Wang, H., McNeill, S.D., Raymond, R.K., Appling, D.R.,

FEATURES	Source
JOURNAL	Shachar-Hill, Y., Bonnett, H. J. and Hanson, A. D.
MEDLINE	Isolation, characterization, and functional expression of cDNAs
PUBMED	encoding NADH-dependent methyltetrahydrofolate reductase from
REFERENCE	higher plants
AUTHORS	2 (bases 1 to 1897)
JOURNAL	Roje, S. and Hanson, A. D.
Source	Direct Submission
Source	Submitted (31-AUG-1999) Horticultural Sciences, University of
Source	Florida, Hall Road, Gainesville, FL 32611-0690, USA
Source	Location/Qualifiers
Source	1. 1897
Source	/organism="Arabidopsis thaliana"
Source	/db_xref="taxon:3702"
Source	1. 1785
Source	/EC_number="1.5.1.20"
Source	/function="catalyzes the reduction of
Source	5,10-methylenetetrahydrofolate (CH2-THF) to
Source	5-methyltetrahydrofolate (CH3-THF)"
Source	/note="NADH-dependent"
Source	/codon_start=1
Source	/product="methyltetrahydrofolate reductase mthfr2"
Source	/protein_id="AAD5788.1"
Source	/db_xref="GI:5911427"
Source	/translation="MKVYDKIOSLADGKSPSEFPPEPTKEDGVNLFPEMDPMVAV
Source	GTTFDDIMWGSAGTDLIDLRKQNVYCVSMHMLCTMPKEKIDHLETRNS
Source	GIQNLALRGDPPEPHGDFKFOVEGEGFCALDVNLRKSTIGYFIVTAGEAPHVD
Source	IECNLQASNEAVOSDLELTKRIDAGADLVLPFDITDIPNDCKRQIGISCPIV
Source	PIQIMINRYRGFLNRGATKRIPEVMALDEIKONEEAVAVGILGTEMCKMLA
Source	HOVSKLHLTYLNMERSALAILMIGMIDESKISRSIPMRPAPNVRTEEDVRPLEFMAN
Source	RKRSYSRTKGMDEPOGRMDSRSASVAGSLSDHSPRRADKIQOEVNPLKSVDE
Source	DIQEKFKELCGNLKSPSWSELDLOPEFTIIEQLIKNSGFLTISQPSVAERSR
Source	DSPTQKSGPVGYYVOKAYALFEFQSKEDKADNVEKKALESITTYAVAKSGDWSNTA
Source	QADVNAVWGVGPAKKEITLOPTIDPASFPMKDEAFETMSRMANLYEADPSRNILE
Source	EKNYSYLVSLVENDYINGDIFAFYADL"
BASE COUNT	524 a 362 c 457 g 354 t
ORIGIN	
Query Match	67.9%; Score 378.8; DB 8; Length 1897;
Best Local Similarity	79.9%; Pred. No. 1.36-85;
Matches 446; Conservative	0; Mismatches 112; Indels 0; Gaps 0;
1	TGATGTGACCCCAATATTTATGATGACGAGATATTCGCAATTTGTGACAGACTGC 60
566	TGATGTGACCTGACGCTTTCATGATGACGATATTTCTCAAGTTGTGATGATGTC 625
61	GCCAAATTTGGAATTAACGTGTCTTATGATGACGATATTCGCAATTTGTGACAG 120
626	GCCAAATTTGGAATTAACGTGTCTTATGATGACGATATTCGCAATTTGTGACAG 685
121	GCTTATTCGCGATGACGAGGTTTGTGCAAAACAAAGATCCAGCGACATATGAGCGT 180
686	GCTTTTGGCTATGATGCTGTTTCTGCAAGACTAAGATACCGATTGAGTATGGCTGCT 745
181	TAGAGCCTATCAAGCAATGAAGAAGCTGTCAAGGCTTATGATGATCAGCTGGAACTG 240
746	TGAGCCTATCAAGCAATGAAGAAGCTGTCAAGGCTTATGATGATCAGCTGGAACTG 805
241	AAATGCTCAAAAAGATTTTACGATCGATGAATTAAGACATTCATCTTATACCTAAATA 300
806	AGATGTGTAAAAAGATGTTGGCTCATGAGTCAAGCTCTTCATCTTATACACTTAAACA 865
301	TGAGAAATTCGATTTGGCAATTAATGAACCTTGAGCTTAAATGAAGACTGCCAAGTTT 360
866	TGAGAAATTCGCTTTCGCAATTAATGAATTAATGAATTCGATGATGATGATGATGATG 925
361	CTAGGTCCTTACCTTGAGAGCGCCGCAAAATGTTTCCGCTTAAAGAGATGTCGGTC 420
926	CTGCTTCTTACCTTGAGAGCGCCGCAAAATGTTTCCGCTTAAAGAGATGTCGGTC 985
421	CAATCTTTTGGGCAATTCAGCAAAAAGCTTCATATCAAGAGCATTGAGATGGATCAAT 480

[illegible]

```
/note="ecotype: Columbia"
1..40
41..1825
/note="putative methylenetetrahydrofolate reductase"
/codon_start=1
/product="At2g44160/F6E13.29"
/protein_id="AAK91450.1"
/db_xref="GI:15125810"
/translation="MKVIDKIOSLADDEKTAFFSEFFPKIEDVDNLFERMDRAVAY
GPFCDITWAGSGSTADLTLDIASRMQNVVCESMMLCTIMPEVEKIDHALETIRSN
GIQWVLAIRGPHQDREYVEGGEDLADLVNIRSKYGDYITVAGYPEAPHDY
IGENGLASNEAYOSDLEYLKKRIDAGADLIVQLFYDIDFLKFNDCROJISCPY
PGIMPINNYRGFLRMTGECRTKIPYEMVAALBEPKIDNEEAVKAYGIHGTCEKKMLA
HGKSLHLYTLMEKSAIILNIGMIDESKISRLPWRPAPVPRKEDVPRIFMAN
RKYSYISRTKWEDEPQGRWGSRSASGALSDFSPRRARDKLQOEWVPLKSE
DIOERFELICLNLSKSPWSELDJQLEPRILINELIVNSKGLTINSQPSVNAERS
DSPTVAGWGPVGYOKAYLEFFCSKEKIDAVEKCALPSITYMAVNKGEQWVNTA
QADVNAVWGWGPVPAKEIIOPTIVDPASEFVWMDAEFTWSRSMANLYPEADPSRNLE
EVKNSYLVLSLVENDYINGIDIFAVFADL"
1826..1917
3'UTR
BASE COUNT      525 a      368 c      460 g      564 t
ORIGIN
Query Match      67.9%; Score 378.8; DB 8; Length 1917;
Best Local Similarity 79.9%; Pred. No. 1.3e-85;
Matches 446; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 1 TGATTGTCACCCCAATTATTTATGATACGGAATATATCTCAAAATTTGTGAACGACTGTC 60
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 606 TGATTGTTACTACGCTTTCTATGACATCGATATATCTCAAGTTTGTGAATGATGACTG 665
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GCCAAATTTGGAATACGTTGCTTATGTCACCTGGAATTAAGCCATTAAATTAATGAAG 120
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 666 GGGAAATTTGGATTAAGTTGTCCTCATTTGTTCTGGAATTAAGCCATTAAATTAATGAAG 725
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GCTTTATCCGCAATGACTGGGTTTGGCAAAACAAGAATACAGCTGACATTATGGCTGCTT 180
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 726 GCTTTTGGCTATGACTGGTTCTTCCGAAGCTAAGTACAGTGTAGGATGGCTGCCCT 785
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 TAGAGCTTATCAAGACATGAAGAAGCTGTCAAGGCTTATGGAATTCACCTGGGACTGTG 240
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 786 TGAGGCTTATCAAGGATTAAGCAAGAGCTGTGAAGCCCTATGTTACCTTTGGAACAG 845
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 AAATGTCAAAAAGATTTTGTGCTCATGGAATTAAGACATTCCTTTATACCTAAATA 300
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 846 AGATGTGTAAAAAGATGTTGGCTCATGAGTCAAGTCTTCTACCTCTACACATTGAACA 905
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 TGGAGAAATCTGCATTGGCAATGCAATGAACTTGCCCTAATTGGAAGATGCCAAAGTTT 360
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 906 TGGAGAAATCTGCTTTCGAATATGATGATGTTGATGATGATGATGATGATGATGATGAT 965
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 CTAGGTCCTTACTTGGAGAGCGCCCTGCAAAATGTTTCCGTGTAAAGAAGATGTCGCTG 420
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 966 CTGCTCTTTTACCTGTGAGAGCGCCCTGCAAAATGTTTCCGTACTAAGGAAGATGTGCGCC 1025
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 CAATCTTTTGGGCAAAATCGACAAAAAGCTATATCAAGACCATAGATGGATGGATCAAT 480
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1026 CCAATTTTCTGTGGCAAAACGTCGAAGCTACATTTCTGAAGCAAGGCGCTGGGAAGATT 1085
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 ACCCAACATGGGCTTGGGGTGATTTCTGTAAATCCATCATATGATGATTAATCTGATTAATC 540
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1086 TCCCAACAGGCGCGTGGGATGATTCACGCGCTTCATATGATGATGATGATGATGATGAT 1145
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 AGTTCAATGGCGGCAAGCG 558
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1146 AGTTTCACGTCGCGGAG 1163
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 3
AF370515 2034 bp mRNA linear PIN 30-APR-2001
LOCUS Arabidopsis thaliana putative methylenetetrahydrofolate reductase
DEFINITION (At2g44160; F6E13.29) mRNA, complete cds.
```

```
ACCESSION AF370515
VERSION AF370515.1 GI:13877628
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2034)
REFERENCE Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,
AUTHORS Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Chauk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sekurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL Direct Submission
TITLE Submitted (17-APR-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Southwick,A.,
Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J.,
Bowser,L., Jones,T., Banh,J., Chen,H., Chauk,R., Chung,M.K.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.
Southwick,A. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.
FEATURES
source
1..2034
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="11"
/clone="RAFL11-04-D03"
/note="this clone is in pb1uescript
ecotype: Columbia"
1..2034
/gene="At2g44160; F6E13.29"
91..1875
/gene="At2g44160; F6E13.29"
/codon_start=1
/product="putative methylenetetrahydrofolate reductase"
/protein_id="AAK3892.1"
/db_xref="GI:13877629"
/translation="MKVIDKIOSLADDEKTAFFSEFFPKIEDVDNLFERMDRAVAY
GPFCDITWAGSGSTADLTLDIASRMQNVVCESMMLCTIMPEVEKIDHALETIRSN
GIQWVLAIRGPHQDREYVEGGEDLADLVNIRSKYGDYITVAGYPEAPHDY
IGENGLASNEAYOSDLEYLKKRIDAGADLIVQLFYDIDFLKFNDCROJISCPY
PGIMPINNYRGFLRMTGECRTKIPYEMVAALBEPKIDNEEAVKAYGIHGTCEKKMLA
HGKSLHLYTLMEKSAIILNIGMIDESKISRLPWRPAPVPRKEDVPRIFMAN
RKYSYISRTKWEDEPQGRWGSRSASGALSDFSPRRARDKLQOEWVPLKSE
DIOERFELICLNLSKSPWSELDJQLEPRILINELIVNSKGLTINSQPSVNAERS
DSPTVAGWGPVGYOKAYLEFFCSKEKIDAVEKCALPSITYMAVNKGEQWVNTA
QADVNAVWGWGPVPAKEIIOPTIVDPASEFVWMDAEFTWSRSMANLYPEADPSRNLE
EVKNSYLVLSLVENDYINGIDIFAVFADL"
BASE COUNT 568 a 386 c 483 g 597 t
ORIGIN
Query Match 67.9%; Score 378.8; DB 8; Length 2034;
Best Local Similarity 79.9%; Pred. No. 1.3e-85;
Matches 446; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

Db 656 TGATTGTACTGAGCTTTTCTATGATCTGATATATTCCTCAAGTTTGTGATGATTGTC 715  
QY GCCAAATTTGGAAATACGTCGCTATTTGACCTGGAATTTGGCCATTTATTTATTTCAAGG 120  
Db 716 GCGAAATTTGGGATTTACTTTGCTCCATTTGCTCGAATTTATGCTTTATTACTACAGAG 775  
QY 121 GCTTTATCCGACGACTGGGTTTTCGCAAAACAAAGATACACGCTGACATTTATGCTGCTT 180  
Db 776 GCTTTTGGGATTTGACTGTTTCTGCAAGACTAAGATACCAATTTAGAGTGATGGCTGCT 835  
QY 181 TAGAGCTTTCAAGGACATGGAAGAGCTGTCAAGGCTTATGGAATTCACCTGGGAACTG 240  
Db 836 TGAGGCTTTCAAGGATTAAGAGAGCTGTGAAAGCCATTTGATATTCACCTTGGAGACAG 895  
QY 241 AATGTGCAAAAAGATTTAGCTTGAATTTAGACATTTGACATTTTATACCTAATAA 300  
Db 896 AGATGTGTAAGAAAGATTTGGCTCATGAGTCAAGTCTTTATCTTCAACATTTGAACA 955  
QY 301 TGGAAATTTGCAATTTGCAATTTGATGATGACCTTGGCTTAATTTGAAGTCCAAAGTTT 360  
Db 956 TGGAGAAATCTGCTCTTGCATATTTGATGATGATGATGATGATGATGATGATGATGAT 1015  
QY 361 CTAAGGCTTTACCTTGGGAGACGCTGCAATGTTTCCGTGTTAAAGAGTTCGCTC 420  
Db 1016 CTGCTTTCTTACCTGAGAGACGCTGCAATGTTTCCGTGTTAAAGAGTTCGCTC 1075  
QY 421 CAATCTTTTGGGCAATTCGACCAAAAGCTACATATCAAGGACCAATGAGATGGATCAT 480  
Db 1076 CCATTTTCTGGGCAAAAGCTTCCAAAGAGCTACATTTCTAAGAACCAAGGCTGGAGATT 1135  
QY 481 ACCCATGTTGGGCTTTGGGCTGATTCCTGTAATCATCATGATGTCATTTGATGATGATG 540  
Db 1136 TCCCAACAAAGCGGTGGGCTGATTCACGAGTGTGATGATGATGATGATGATGATGATG 1195  
QY 541 AGTTATGCGGCGACGTG 558  
Db 1196 AGTTCTCACGTCCGCGAG 1213

RESULT 4  
ATH245414 2037 bp mRNA linear PLN 01-SEP-1999  
LOCUS Arabidopsis thaliana mRNA for methylenetetrahydrofolate reductase.  
DEFINITION AJ245414.1 GI:5823582  
ACCESSION 5,10-methylenetetrahydrofolate reductase; methylenetetrahydrofolate  
VERSION reductase; MTHFR1 gene.  
KEYWORDS Arabidopsis thaliana.  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 2037)  
AUTHORS Ravanell, S., Rebelle, F. and Douce, R.  
TITLE Folate metabolism in higher plants: cloning of a cDNA for  
5,10-methylenetetrahydrofolate reductase in Arabidopsis thaliana  
JOURNAL Unpublished  
AUTHORS Ravanell, S.  
TITLE Direct Submission  
JOURNAL Submitted (27-JUL-1999) Ravanell S., DBMS / PCV, URA-CNRS 576,  
CEA-Grenoble, 17, rue des Martyrs, 38054 Grenoble Cedex 9, FRANCE  
FEATURES  
source  
1. 2037  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
gene 68..1852  
/gene="MTHFR1"  
CDS 68..1852  
/gene="MTHFR1"  
/EC\_number="1.5.1.20"  
/function="reduction of methylenetetrahydrofolate to  
methylenetetrahydrofolate"

/codon\_start=1  
/product="methylenetetrahydrofolate reductase"  
/protein\_id="CAB53783.1"  
/db\_xref="GI:5823583"  
/translation="MKVLDKIQSLADEGKTAFSEFFPPKTEGDVQDLNFERMDRVAY  
GPFCDITWAGSGSTADLTLDIASRMQSYLVDSMMHLCTNMPVKKIDHAELETISN  
GIONVALRDPDHPGDDKFPVQGGFDCALDVNHRKSYGDYFGITTVAGYEAPDDV  
IGBNGLASNEAYOSDL EYLKKRIDAGADLIYNQLFPYDPIFLKFPVNDQDQISCPY  
PGIMPINNRYGFLRMTEGCTKLPVYMAALEPIKONEAVATGIGHGTEKRMILA  
HGKSLHLTYLNLKESALALIMNLKIDSKISRSULPMRPANVPETKEDVRPIFWAN  
RPSYSISRTKGWDEPQGRWDSRSASYSALSDHSPRDRADKLODEWVPLKVE  
DIOEKFEKLIDGNLKSMPSELDDGLPETRIINDEOLIKVNSGFLTINSQVNSRSN  
DSPTVMGPGVGVYXKAYLEFCSKEKIDAVYEKALPSITLYAVNVEGEQWVSVT  
QADVNAVTVGVPFAKELIOPITVDYDIPASFPVWMDENAEETMSRMANLYPEADSRNMLE  
EVANSYLYVLSVENDYINDIRAVPDL"

BASE COUNT 583 a 385 c 480 g 589 t  
ORIGIN  
Query Match 67.6%; Score 377.2; DB 8; Length 2037;  
Best Local Similarity 79.7%; Pred. No. 3,2e-85;  
Matches 445; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 1 TGATTGTACCCAAATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATG 60  
Db 633 TGATTGTACTACACCTTTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 692  
QY 61 GCCAAATTTGGAAATACGTCGCTATTTGACCTGGAATTTGGCCATTTATTTATTTCAAGG 120  
Db 693 GCGAAATTTGGGATTTACTTTGCTCCATTTGCTCGAATTTATGCTTTATTACTACAGAG 752  
QY 121 GCTTTATCCGACGACTGGGTTTTCGCAAAACAAAGATACACGCTGACATTTATGCTGCTT 180  
Db 753 GCTTTCTTACCTGAGAGACGCTGCAATGTTTCCGTGTTAAAGAGTTCGCTC 812  
QY 181 TAGAGCTTTCAAGGACATGGAAGAGCTGTCAAGGCTTATGGAATTCACCTGGGAACTG 240  
Db 813 TGGAGCTTATCAAGGATTAAGAGAGCTGTGAAAGCCATTTGATGATGATGATGATGATG 872  
QY 241 AATGTGCAAAAAGATTTAGCTTGAATTTAGACATTTTATACCTAATAA 300  
Db 873 AGATGTGTAAGAAAGATTTGGCTCATGAGTCAAGTCTTTATCTTCAACATTTGAACA 932  
QY 301 TGGAAATTTGCAATTTGCAATTTGATGATGACCTTGGCTTAATTTGAAGTCCAAAGTTT 360  
Db 933 TGGAGAAATCTGCTCTTGCATATTTGATGATGATGATGATGATGATGATGATGATGAT 992  
QY 361 CTAAGTCTTTACCTTGGGAGACGCTGCAATGTTTCCGTGTTAAAGAGATGTCGCTC 420  
Db 993 CTGCTTTCTTACCTGAGAGACGCTGCAATGTTTCCGTGTTAAAGAGATGTCGCTC 1052  
QY 421 CAATCTTTTGGGCAAAATTCGACCAAAAGCTACATATCAAGGACCAATGAGATGGATCAT 480  
Db 1053 CCATTTTCTGGGCAAAAGCTTCCAAAGAGCTACATTTCTAAGAACCAAGGCTGGAGAACT 1112  
QY 481 ACCCATGTTGGGCTTTGGGCTGATTCCTGTAATCATCATGATGTCGATTAATCTGATTATC 540  
Db 1113 TCCCAACAAAGCGGTGGGCTGATTCACGAGTGTGATGATGATGATGATGATGATGATG 1172  
QY 541 AGTTATGCGGCGACGTG 558  
Db 1173 AGTTCTCACGTCCGCGAG 1190

RESULT 5  
AY122922 1810 bp mRNA linear PLN 03-JUL-2002  
LOCUS Arabidopsis thaliana putative methylenetetrahydrofolate reductase  
DEFINITION MTHFR1 (At3g59970) mRNA, complete cds.  
ACCESSION AY122922  
VERSION AY122922.1 GI:21689666  
KEYWORDS FLI\_cDNA.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana

REFERENCE  
AUTHORS

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae, eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1810)

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,Y.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yu,G., Bower,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Shinozaki,K., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Arabidopsis Open Reading Frame (ORF) Clones

TITLE  
JOURNAL  
COMMENT

2 (bases 1 to 1810)  
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,Y.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yu,G., Bower,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Shinozaki,K., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.  
Submitted (17-JUN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.

The Saik, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,Y.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yu,G., Bower,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Miranda,M., Nguyen,M., Palm,C.J., Shin,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

## Location/Qualifiers

1. 1810  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="U18407"  
/note="This clone is in PUNI 51.  
ecotype: Columbia"  
1. 1810  
/gene="At3g59970"  
1. 1779  
/gene="At3g59970"  
/codon\_start=1  
/evidence=experimental  
/product="putative methylentetrahydrofolate reductase MTHFR1"

gene  
CDS

/protein\_id="AA067455.1"

/db\_xref="GI:2168967"  
/translation="MKVVDKIKSVTEGQTAFFSEFPFKTEGVENLFEKMDRLVY  
GPTFCDIWAGAGSTADLTLETASRMQVNICVETMHCTNMPEKIDHAELETIRSN  
GIONVIALRGDPHGDGKFEVGGGFCALDVNIIRSKYDGYGVYGAPEAHDPV  
TEADGLATPEYOSDLAYLKKRKNVAGADLYTOLPYDPIFLKPVNOCROGICNPY  
PGMPSINVKGLRMAGCECKTIRPRLTALAEPIKNDENAYKATGIRHATMCKKILA  
HGITSLLTTLNWDKSAIGILMNLGLIDESKISRLEPWRPRANVFRKEDVRLFWAN  
RKSYISRTKWNDDPHGRWDSHAAYSTLSDYQFAPRKGRDKLQDEVVPLKSI  
DVEQEKELCIGNLSPWSEIDGQPEPKTILINQIGKINSNGEFLTINSQSVAAKS  
DSPALGMGPGGYVOKAYLEFGSKDKLDLTVKSKAFPSITTYMAVVKSNWVNSG  
ESDVNAVTVGWVPAKEVIOPTIYDPAKFKWDEAFETWSSMANLYPEDPSKILLE  
EYKNSYIYSLVDNNYINGDIFSVA"

1780. 1810

/gene="At3g59970"

508 a 372 c 427 g 503 t

3' UTR

BASE COUNT

## ORIGIN

Query Match 64.3%; Score 358.8; DB 8; Length 1810;  
Best Local Similarity 78.0%; Pred. No. 1.5e-80;  
Matches 432; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 1 TGATTGCAACCAATTTATTATGATACGATATATCCCAATTTGCAAGACGTC 60  
DB 566 TGAATGTGACACCTTTCTATGATCTGATATATTTCTCAAGTTTGTAAGATGTC 625  
QY 61 GCCAAATGGAATTAACGTCCTATTTACCTGGAAATTAATTAATTAACAG 120  
DB 626 GGCAATCGGATTAATTTGTTCCCATTTGTTCTGGAATTAATTAATTAACAG 685  
QY 121 GCTTATCCGATGACCTGGTTTGGCAAAACAAAGATTAACGCTGACATTAAGCTGCTT 180  
DB 686 GGTTCCTGCGATGCTGGCTGTTCTGTAAGACCAAGATTAACCGGACCTGACCTGCT 745  
QY 181 TAGAGCCATGACGACAAAGAAAGACCTGCAAGCTTATGAAATTCACCTGGAACTG 240  
DB 746 TAGAGCCATTAAGGATTAATGACGAGGCTGTTAAGCCTTAATGAAATTCACCTGGAACTG 805  
QY 241 AATGTGCAAAAGATTTTACCTCATGGAATTAAGACATTCATTTATACATAATA 300  
DB 806 AATGTGCAAAAGATTTTGGCCCATGGAATTAATTCATTCCTCATACATAATAAG 865  
QY 301 TGAAGAAATCTGCAATGGAATTAATGAACCTTGCCCTTAATGAAGAGTCCAAAGTTT 360  
DB 866 TGAAGAAATCTGCAATGGAATTAATGAACCTTGCTGATGATGAGTCAAAATTTT 925  
QY 361 CTAGGCTTACCTTGAGAGCCCTGCAAAATGTTTCCGTTAAAGAGATGTCCTG 420  
DB 926 CTCGTTCTCTACCTTGAGAGCCCTGCAAAATGTTTCCGTTAAAGAGATGTCCTG 985  
QY 421 CAATCTTTGGGCAATGCAACCAAAAGCTACATATCAAGACCATATGATGGATCAAT 480  
DB 986 CAATTTCTGGGCAACCGCTCAAAAGCTACATATCAAGACCATATGATGGATCAAT 1045  
QY 481 ACCACATGAGGCTGTTGGGCTATTCCTGTAATCATATATGATGATATGATATTC 540  
DB 1046 TCCACATGAGGCTGTTGGGCTATTCACACAGTACATGATGATGATGATATTC 1105  
QY 541 AGTTCATGCGCCA 554  
DB 1106 AGTTTGGCGCCA 1119  
RESULT 6  
AF181966  
LOCUS  
DEFINITION  
Arabidopsis thaliana methylentetrahydrofolate reductase MTHFR1  
ACCESSION  
AF181966  
VERSION  
AF181966.1 GI:5911424  
KEYWORDS  
ORIGIN  
Arabidopsis thaliana.  
Arabidopsis thaliana  
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1994)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (31-AUG-1999) Horticultural Sciences, University of











```

repeat_region      complement(5593..5613)
                    /rpl_family="At_rich"
repeat_region      complement(6730..6733)
                    /rpl_family="At_rich"
gene                7192..8658
                    /gene="At2g43890"
                    /note="FE13.2"
mRNA                join(-7192..7632,7743..8032,8122..8350,8440..8658)
                    /gene="At2g43890"
CDS                 join(1192..7632,7743..8032,8122..8350,8440..8658)
                    /gene="At2g43890"

```

```

/protein_id="AAC23398.1"
/db_xref="GI:3212847"
/translation="MDNKKLVAVILMFSSFLMKSTAASNVYVSGAPDQRTDS
TKFLFAMQACRSAAYTVVPRGSFLNPVFRPDCSRITFEIYGTVIAPSDYSGK
LNGSGWILFVAVNRISLIGTLDARCASTACRKGKSCPPGARSMTFNANVYVS

```

GLISINSQTTTLVINSNCNVIARRKVLVADDDSPNIDHGLVGSAGVAVITDGLTHTDGLD  
DCISIPGRNLTLYMSKLCNCPGHGHSIGSPANAGENTLILNFSFSDNMGVRL  
KTVARSTFVAVNVLFOÜLIMKNÖNPILYDÖNCCSNOCCKÖSSGKISÖVYTRNL  
ÖGTSRTQALTFDCRSRNPQÖAIRLHDIKTLFENGSRATSTCKNIGKAVGVMMPOGCL  
H

gene	complement(9101..14090)
gene	/gene="At2g43900"
mrna	/note="FEI3.3"
complement(10101..9552,10049..10409,10500..10623,10717..10984,11093..11133,11219..11384,11587..11671)	

CDS	11768..11895,12004..12811,12910..13162,13359..14090)))
	/gene="Ar2g4390"
	Complement(join(9101..9952,10049..10409,10500..10623,
	10717..10894,11093..11133,11210..11384,11587..11671),
	11768..11895,12004..12811,12910..13162,13359..14090))

```
/gene="At2g43900"  
/product="putative inositol polypophosphate 5'-phosphatase"  
/protein_id="AAC23399.1"
```

```

/obj_xref="G1:3212848"
/translation="MDIINNRRDENDDEEALSMSSVPPPKIHSHSHQLRATGG
KGNHRQRHSLDDIPKITEIVSGGICGSDSDDEFPYPIATTTNSSPPPTGGDGDST
DYLHGEIEDPQPLPEFVGSGGGMFKVPKPSLHSPAPCLDELRHPHKLQVQVH
FYFNIACTFETQIMAGCGSVPRFWMNDAPKPGQISGRVORGLDAPAPORSAISNPP

```

[illegible]

LADSGMIWGCNGNGLIOMDNGRRLDRFHIOCAVCEPCTGERTGRTGYGVYSGXJAI  
 DLGNGIACGWANNAVYKIMAAADGYIFSLATGAGIKGAWYISPGJLGLIIRSELA  
 ERTYAOTDSYRLLITGSMWVGOGKASHDLMSWLGSAVSYGLIIVGLDVEVEMGAGFLA  
 MSAAKSVGNGNCTIGYVITGKTRDLAEVFRGSGYQIAGLILSWAKKATDTH  
 VGLIDVAAPVCGEGRATCNKSGVGLRTVFRDRCINCHIAALEAVNVRRAADPDH

YKIMSTPSSMHNHPAAGVSTGSHITTKSANNANVTBETKODLEADMYVPFGDENY  
RLGIGISDEARDFVSORSDFWJREDDLSREBMAGRFGOMARPAITPEPYKERSH  
PGAGVDSGKKRIPAMCDRIYFDRTRSESGDCPCVASIMUYACMDVESHEN  
KPRACEFHAKIEHVDRSVRQDEGRITKTEKRALINDLRNYPETIYSSNSIYLQND

[illegible]

```
repeat_region DSSSKSHKSSGEGDSSLSLTKRMELIAHNSRRKGRVTYROL"
               14103..14213
               /rpt_family="(GAGAA)n"
gene complement(14748..17292)
```

/gene="ALG43910"  
 /note="F6E13.4; supported by full length cDNA:  
 Ceres:24370"  
 complement(join<14748..14963,15285..15363,15476..155222  
 15617..15716,15874..15974,16059..16145,16677..16807,

```

CDS
/gene="At2g43910"
complement(join(14925..14963,15285..15363,15476..15522,
15617..15716,15874..15974,16059..16145,16677..16807,
17119..17218))
/gene="At2g43910"
/codon_start=1
/product="putative methyl chloride transferase"
/protein_id="AAC23400.1"
/db_xref="GI:3212849"
/translation="MAEQNSDOSNGSNVLPPEEVAFLHKTVEEGWKEGMEIEI
TPWDGRATPLILVHVDSSILPLGRALVPGGGHDVVAASPERFVGLDISESLA
KANETYGSSPKAEYFSFKEDVFTMRPELDELFDVDFPCALIEPMRPMASMEEL
LKPDGLITLMTPTDHYGEPPIKVDSTFEVLPVIGFAVSEENPHAIPIPRKKE
KLGRRKKIN"
complement(15108..15141)
/rpt_family="AT_rich"
complement(16570..16608)
/rpt_family="AT_rich"
complement(17642..17687)
/rpt_family="AT_rich"
17774..17794
/rpt_family="AT_rich"
17926..17970
/rpt_family="AT_rich"
18152..18196
/rpt_family="AT_rich"
complement(19342..21515)
/gene="At2g43920"
/feature="FEI3.5"
complement(join(<19342..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,
21416..>21515))
/gene="At2g43920"
complement(join(19586..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,
21416..21515))
/gene="At2g43920"
/codon_start=1
/product="unknown protein"
/protein_id="AAC23401.1"
/db_xref="GI:3212850"
/translation="MAEQNSDOSNGSNVLPPEEATFQPOVAEGGMDKCEDCV
TPWDGRATPLILHLDSSALPLGRIVPGGGHDVVAASPERFVGLDISKALN
KANETYGSSPKAEYFSFKEDVFTMRPELDELFDVDFPCALIEPMRPMASMEEL
LKPDGLITLMTPTDHYGEPPIKVDSTFEVLPVIGFAVSEENPHAIPIPRKKE
KLGRRKKIN"
23418..25849
/gene="At2g43930"
/feature="FEI3.6; contains a protein kinase domain profile
(PPOC00100)"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..>25849)
/gene="At2g43930"
join(23455..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
/codon_start=1
/product="putative protein kinase"
/protein_id="AAC23427.1"
/db_xref="GI:3212876"
/translation="MDHRRRLMTIKILINSSSTLLFEKIVPILHREGSDRALPYT
PFFCSICRNALGHGNTSEPHIVLLQCRKSKMCFASPEEHKLGKVIDEENHRK
GKRGVLAAREEVQOLKNSHPIYVIGTRESINSIMQFVPGGSISSLAKFGS
FPEKHNRRLDNLATISHKSYKNIHQOSFKIEDPRKKKOS"
complement(27269..30414)
/gene="At2g43940"
/feature="FEI3.7"
complement(join(<27269..27361,27641..27771,27853..27984,
28336..28397,28479..28537,28635..28784,28865..28981,
29149..29343,30043..>30414))
/gene="At2g43940"
complement(join(27269..27361,27641..27771,27853..27984,

```

```

Query Match      21.8%; Score 121.8; DB 8; Length 109741;
Best Local Similarity 79.6%; Pred. No. 1.4e-20;
Matches 144; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 149 AACAAAGATACCGCTGACATTTATGCTGCTTTAGAGCCTATCAAGACATGAAGAGC 208
DB 93752 AATGTGATACACAGTTGAGGTATGCTGCTTGAGCCTATCAAGATACGAAGAGC 93811
OY 209 TGTCAAGCCTTATGCAATTTACCTGGGAACATGCAAAATGTCAAAAGATTTAGCTCATG 268
DB 93812 TGTGAAGCCTATGATATTCACCTTGGAACAGATGTGTAAGAAAGTTGGCTCATGG 93871
OY 269 AATTAGACATTCATCTTTATACACTAAATATGAGAAATCTCATTTGCAATATCAT 328
DB 93872 AGTCAAGCTCTTCATCTTACACATTCACATGAGAAATCTGCTTGCATATTCAT 93931
OY 329 G 329
DB 93932 G 93932

RESULT 13
SPAC56F8 37093 bp DNA linear PLN 28-JUN-2000
LOCUS S.pombe chromosome I cosmid c56F8.
DEFINITION 269728
ACCESSION 269728
VERSION 269728.1 GI:1204222
KEYWORDS alpha-glucosidase; escl; methylenetetrahydrofolate reductase;
microsomal signal peptidase;
para-hydroxybenzoate--polyprenyltransferase; probable membrane
protein.
SOURCE fission yeast.
ORGANISM Schizosaccharomyces pombe
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomycetes.
1 (bases 1 to 37093)
Walsh,S.V.
Pearson,D.; Churcher,C.M.; Barrell,B.G.; Rajandream,M.A. and
DIRECT SUBMISSION
Submitted (14-FEB-1996) Schizosaccharomycetes pombe chromosome I
sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge
CB10 1RQ E-mail: barrell@sanger.ac.uk
Notes:
details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S_pombe)
Protein coding regions (CDS) have been predicted with the help of
computer analysis using the Genefinder program in Pombase (an ACEDB
database) with additional predictions for the branch-acceptor sites
supplied by the program Sp3splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.
CDS are numbered using the following system eg SPAC5H10.01c, SP (S.
pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
(complementary strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
Cosmid c56F8 is overlapped at the 5' end by cosmid c30D11 and at
the 3' end by cosmid c22A12.
Location/Qualifiers
1..37093
/organism="Schizosaccharomycetes pombe"
/strain="972h-"
/db_xref="taxon:4896"
/chromosome="I"
/map="IL"
FEATURES
source

```





```

Db 674 TCATCATCAGCAGCTTTCTTGGAGCTGCACATCTTCCGCTTTGTAAGCATGCA 733
QY 61 GCCAATTGGAAATTAAGTGTCTTATTTACTGGAAATTAATTAATTAATTAACAGG 120
Db 734 CCGACATGGGACATCACTTCCCATTCGCCCCGATCTTCCATCCAGGCTACCACT 793
QY 121 GCTTATCCGATGACTGGGTTTTCAGAAACAAAGATACCAAGCTGACATTAATGGCTGCTT 180
Db 794 CCTTGGGAGCTTGTGAAGCTGTCCAGCTGAGGTGCCACAGAGATCAAGAGCTGA 853
QY 181 TAAGGCTTATCAAGAGCAATGAAGCTGTCAAGGCTTATGAATTCACCTGGAGCTG 240
Db 854 TTGAGCCAAATCAAGACAAAGATGCTGCCATCCGAACTATGCAATCGAGCTGGCCGTGA 913
QY 241 AAATGTGCAAAAAGATTTTATGCTCATGGAATTAAGACAT--TGATCTTTATACACTAA 297
Db 914 GCTGTGCCAGAGCTTGTGGCCAGTGGCTTGTGCCAGGCTTCACATCTTACACCTCA 973
QY 298 ATATGGGAAATCTGCATTTGGCAATTAATGAACCTTGGGCTAATTGAAGAGTCCAAAG 357
Db 974 ACCGCGATGGCTACCAAGAGTGTGAAGCGCTGGGATGTGACTGAGAGACCCA 1033
QY 358 TTCTAGGCTCTTACCTTGGAGAGCGCCTGCAGAAATGTTTCCGTGTTAAGAAATGTCC 417
Db 1034 GGCCTC--CCCTACCTTGGCTCTCAGTGGCCACCCCAAGCGCGAGAGAGATGTAC 1090
QY 418 GTCCAACTTTTGGCAATTCGACCAAAAGCTACATATCAAGAGCCATAGATGGATC 477
Db 1091 GTCCCAATCTTGTGGGCTTCAGCAAGAGTTACATCTACCTTACCCAGAGTGGAGC 1150
QY 478 AATACCAATGGGCTTGGGCTGATTCCTGTAATCCATCATATAGTGCATTAATCTGATT 537
Db 1151 AGTCCCTACAGGCGCTGGGCAATCTCTTCCCTTGGGAGCTGAAGGACT 1210
QY 538 ATCAGTTCAT 547
Db 1211 ACTACTCTT 1220

```

Search completed: February 14, 2003, 21:20:23  
 Job time : 3017 secs

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_



XX Falco SC, Fomodu LO;  
PI  
XX  
DR WPI: 2000-182429/16.  
DR P-PSDB; ANY44740.  
XX

PT Novel tetrahydrofolate metabolism enzyme used to alter the level of  
PT tetrahydrofolate metabolism in plants and seeds -  
XX  
PS Claim 3: Page 32; 37pp: English.  
XX

CC The present sequence is a cDNA clone sfl1\_pK0017.d12 encoding  
CC 5,10-methylenetetrahydrofolate reductase (MTHFR). The clone was  
CC isolated from a sfl1 cDNA library which was prepared using soybean  
CC immature flowers. MTHFR plays a role in the  
CC synthesis of methionine. The present sequence is used in the  
CC construction of a chimeric gene to alter the level of tetrahydrofolate  
CC metabolism enzymes in plants. The enzyme may provide target to facilitate  
CC design and/or identification of inhibitors that may be useful as  
CC herbicides. The polynucleotide is also useful as a source of probes for  
CC genetically and physically mapping the genes and as markers for traits  
CC linked to the genes.  
XX

SQ Sequence 558 BP; 170 A; 109 C; 116 G; 163 T; 0 other;

Query Match 100.0%; Score 558; DB 21; Length 558;  
Best Local Similarity 100.0%; Pred. No. 9.1e-156;  
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGATGTCACCAATATATTTATGATACGATATATCTCATAATTTGTGACGATGTC 60  
DB 1 TGATGTCACCAATATATTTATGATACGATATATCTCATAATTTGTGACGATGTC 60  
QY 61 GCCAAATGGATACGATGTCATTTGTACCTGGAATTTAGCCCTTAATATACAGG 120  
DB 61 GCCAAATGGATACGATGTCATTTGTACCTGGAATTTAGCCCTTAATATACAGG 120  
QY 121 GCTTATCCGATGATGCTGTTTGGCAAAACAAAGATACAGCTGATATGCTGCTT 180  
DB 121 GCTTATCCGATGATGCTGTTTGGCAAAACAAAGATACAGCTGATATGCTGCTT 180  
QY 181 TAGAGCCTATCAAGCAATGATGAGAGCTGTCAAGCCTATGATTCACCTGGAGCTG 240  
DB 181 TAGAGCCTATCAAGCAATGATGAGAGCTGTCAAGCCTATGATTCACCTGGAGCTG 240  
QY 241 AAATGTGCAAAAGATTTTACCTCATGATTAAGACATTCATTTATACATAATA 300  
DB 241 AAATGTGCAAAAGATTTTACCTCATGATTAAGACATTCATTTATACATAATA 300  
QY 301 TGGAGAAATCTGATTTGCAATGATACCTTGCCCTAATTTGAAGAGTCAAAAGTTT 360  
DB 301 TGGAGAAATCTGATTTGCAATGATACCTTGCCCTAATTTGAAGAGTCAAAAGTTT 360  
QY 361 CTAGTCTCTTACCTTGAGAGCGCCCTGCAAAATGTTTCCGCTTAANAAGATTCCTGTC 420  
DB 361 CTAGTCTCTTACCTTGAGAGCGCCCTGCAAAATGTTTCCGCTTAANAAGATTCCTGTC 420  
QY 421 CAATCTTTTGGGCAATGACAAAAAGCTACATATCAAGGACCATAGATGGATCAAT 480  
DB 421 CAATCTTTTGGGCAATGACAAAAAGCTACATATCAAGGACCATAGATGGATCAAT 480  
QY 481 ACCGCAATGGGCTTGGGGGATTCCTGTAATCCATCAATGAGGACATATTCATATATC 540  
DB 481 ACCGCAATGGGCTTGGGGGATTCCTGTAATCCATCAATGAGGACATATTCATATATC 540  
QY 541 AGTTCATGCGGCCACGTG 558  
DB 541 AGTTCATGCGGCCACGTG 558

RESULT 2  
AAC43311  
ID AAC43311 standard; DNA; 1821 BP.

XX AAC43311;  
AC  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38803.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
PD  
XX  
PD 06-SEP-2000.  
XX  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126268.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140699.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142054.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142927.  
PR 13-JUL-1999; 99US-0143547.  
PR 14-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144062.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148177.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154039.  
PR 16-SEP-1999; 99US-0154018.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 28-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 67.9%; Score 378.8; DB 21; Length 1821;  
Best Local Similarity 79.9%; Pred. No. 2.5e-102;  
Matches 446; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
  
QY 1 TGATGTGACCCCAATTTATATGATACGATATATCTCTCAAAATTTGTGAACACGTC 60  
DB 602 TGATGTGATCTGCTTTCTATGATCTGATATATCTCAAGTTTGTGATGATGTC 661  
  
QY 61 GCCAAATGGAATTAACGTCCTCTATGTCTACCTGGAATTAAGCCATTAATATTAAGG 120  
DB 662 GCCAAATGGAATTAAGTTGTCCTCATTTGTCCTGGAATTAATTAATTAACAGAG 721  
  
QY 121 GCCTTATCCGATGACTGGGTTTGCACAAACAAAGATACAGCTACATTAATGGCTGCTT 180  
DB 722 GCCTTTGGCTATGACTGGTTTCTGCAAGACTAAGATACAGTTGATGATGGCTGCTT 781  
  
QY 181 TAAGGCTATCAAGACATTAAGAAAGCTGTCAAGCTTATGAAATTACCTGGAGACTG 240  
DB 782 TGAGGCTATCAAGATTAACAGAAAGAGCTGTGAAGCCCTATGATTAACCTTGAGACG 841

```
OY 241 AATGTCGCAAAAGATTTAGCTCATGGAATTAGACATTGCATCTTATTACACTAATA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 842 AGATGTGTAAAGATGTTGGCTCATGAGTCAAGTCTCTCATCTACATGAGACA 901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 301 TGSAGAAATCTGCATTGGCAATACATACATGAGCTTGCCCTAATTGGAAGTCCAAAGTT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 902 TGSAGAAATCTGCCTTGCATATTTGATGAATCTTGGTATGATGTAGTGCACAAATTT 961
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 361 CTAGTCTTACCTTGGAGAGCCCTGCAAGTGTTCCTGCTGCTTAAGAAGATGCCGTC 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 962 CTCGTTCTTTACCTTGAGAGCCCTGCAAGTGTTCCTGCTTACAGAGAAATGTGCCGC 1021
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 421 CAATCTTTTGGGCAAAATGCACCAAAAGCTACATATTCAGAGACCATTAGATGGATCAAT 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1022 CCAATTTCTGGGCAAAACCGTCCAAAGAGCTCATTTCTAGAACCAAGGCGTGGGAAGATT 1081
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 481 ACCCAATGGGCGTGGGGTGATTCCTGTAACTCATATGATGTCATTCGATTCATTC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1082 TCCACAAAGCCGCTGGGGTGATTCACGACAGTCTTATATGATGTCACCTCGGATCATC 1141
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 541 AGTTCATGCGGCGCAAGCTG 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1142 AGTTCATGCTCCGCGAG 1159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
AAC49168
ID AAC49168 standard; DNA; 1942 BP.
XX
AC AAC49168;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SPQ ID NO: 60181.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130047.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135628.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
```

PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 04-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 05-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 06-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 09-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 10-AUG-1999; 99US-0147935.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0148175.  
 PR 17-AUG-1999; 99US-0148426.  
 PR 18-AUG-1999; 99US-0148722.  
 PR 20-AUG-1999; 99US-0148723.  
 PR 20-AUG-1999; 99US-0149020.  
 PR 23-AUG-1999; 99US-0149802.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150889.  
 PR 27-AUG-1999; 99US-0151086.  
 PR 27-AUG-1999; 99US-0151087.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151348.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159337.  
 PR 14-OCT-1999; 99US-0159636.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161922.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.  
 Query Match 67.9%; Score 378.8; Db 21; Length 1942;  
 Best Local Similarity 79.9%; Pred. No. 2,6e-102;  
 Matches 446; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
 OY 1 TGAATGTCACCAATATATTTATGATACGATATATTCCTCAATTTTGACACAGTCGTC 60  
 Db 566 TGAATGTCACCAATATATTTATGATACGATATATTCCTCAATTTTGACACAGTCGTC 625  
 OY 61 GCCAATTTGGGATTAAGTTGTCATGATACGATATATTCCTCAATTTTGACACAGTCGTC 120  
 Db 626 GCCAATTTGGGATTAAGTTGTCATGATACGATATATTCCTCAATTTTGACACAGTCGTC 685  
 OY 121 GCTTATCCGATGATGCTGGTTTGCAAAACAAAGATACAGCTGACATTTAGCTGCTT 180  
 Db 686 GCTTATCCGATGATGCTGGTTTGCAAAACAAAGATACAGCTGACATTTAGCTGCTT 745  
 OY 181 TGAAGCTATCAGCAATGAAAGAGCTGCAAGCTTGTGAATTCACCTGGAGACGTC 240  
 Db 746 TGAAGCTATCAGCAATGAAAGAGCTGCAAGCTTGTGAATTCACCTGGAGACGTC 805  
 OY 241 AATATGCAAAAATATTTTACATGATATTAAGATTCATGATTTTACATTAATA 300  
 Db 806 AATATGCAAAAATATTTTACATGATATTAAGATTCATGATTTTACATTAATA 365  
 OY 301 TGGAGAAATCTGCTTGGCAATCTAATGACATTTGGCTTAATGAGAGTCGCAATTT 360  
 Db 866 TGGAGAAATCTGCTTGGCAATCTAATGACATTTGGCTTAATGAGAGTCGCAATTT 925  
 OY 361 CTAGTCTTAACTTGGAGAGCCCTGCAATGTTTCCGTTAAAGAGTCGCTC 420  
 Db 926 CTAGTCTTAACTTGGAGAGCCCTGCAATGTTTCCGTTAAAGAGTCGCTC 985  
 OY 421 CAATCTTTTGGCAATTCGACCAAAAAGCTACATATCAAGGACCATAGAGTCGCAAT 480  
 Db 986 CAATCTTTTGGCAATTCGACCAAAAAGCTACATATCAAGGACCATAGAGTCGCAAT 1045  
 OY 481 ACCACATGCGGCTTGGGATTCCTGTAATCCATATGATGTCGATTAATCTGATATTC 540  
 Db 1046 ACCACATGCGGCTTGGGATTCCTGTAATCCATATGATGTCGATTAATCTGATATTC 1105  
 OY 541 AGTTACGCGGCGACGTCG 558  
 Db 1106 AGTTACGCGGCGACGTCG 1123  
 RESULT 4  
 ID AAC44829 standard; DNA; 1493 BP.  
 AC AAC44829;  
 DT 18-OCT-2000 (first entry)  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 44287.  
 XX  
 XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 RW metabolic pathway; promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 OS  
 XX  
 XX EPI033405-A2.  
 XX

PD 06-SEP-2000.  
XX 25-FEB-2000. 2000EP-0301439.  
XX 25-FEB-1999. 9905-0121825.  
PR 05-MAR-1999. 9905-0123160.  
PR 09-MAR-1999. 9905-0123548.  
PR 23-MAR-1999. 9905-0125788.  
PR 25-MAR-1999. 9905-0126252.  
PR 29-MAR-1999. 9905-0126785.  
PR 01-APR-1999. 9905-0127462.  
PR 06-APR-1999. 9905-0128224.  
PR 08-APR-1999. 9905-0128714.  
PR 16-APR-1999. 9905-0129845.  
PR 19-APR-1999. 9905-0130077.  
PR 21-APR-1999. 9905-0130449.  
PR 23-APR-1999. 9905-0130510.  
PR 23-APR-1999. 9905-0130891.  
PR 28-APR-1999. 9905-0131449.  
PR 30-APR-1999. 9905-0132048.  
PR 30-APR-1999. 9905-0132407.  
PR 04-MAY-1999. 9905-0132484.  
PR 05-MAY-1999. 9905-0132485.  
PR 06-MAY-1999. 9905-0132486.  
PR 07-MAY-1999. 9905-0132487.  
PR 11-MAY-1999. 9905-0134256.  
PR 14-MAY-1999. 9905-0134218.  
PR 14-MAY-1999. 9905-0134219.  
PR 14-MAY-1999. 9905-0134231.  
PR 14-MAY-1999. 9905-0134231.  
PR 18-MAY-1999. 9905-0134370.  
PR 19-MAY-1999. 9905-0134768.  
PR 20-MAY-1999. 9905-0134941.  
PR 21-MAY-1999. 9905-0135124.  
PR 21-MAY-1999. 9905-0135353.  
PR 24-MAY-1999. 9905-0135639.  
PR 25-MAY-1999. 9905-0136071.  
PR 27-MAY-1999. 9905-0136392.  
PR 28-MAY-1999. 9905-0136722.  
PR 01-JUN-1999. 9905-0137242.  
PR 03-JUN-1999. 9905-0137358.  
PR 04-JUN-1999. 9905-0137502.  
PR 07-JUN-1999. 9905-0137754.  
PR 08-JUN-1999. 9905-0138094.  
PR 10-JUN-1999. 9905-0138540.  
PR 10-JUN-1999. 9905-0138847.  
PR 14-JUN-1999. 9905-0139119.  
PR 16-JUN-1999. 9905-0139452.  
PR 16-JUN-1999. 9905-0139453.  
PR 17-JUN-1999. 9905-0139452.  
PR 18-JUN-1999. 9905-0139454.  
PR 18-JUN-1999. 9905-0139455.  
PR 18-JUN-1999. 9905-0139456.  
PR 18-JUN-1999. 9905-0139457.  
PR 18-JUN-1999. 9905-0139458.  
PR 18-JUN-1999. 9905-0139459.  
PR 18-JUN-1999. 9905-0139460.  
PR 18-JUN-1999. 9905-0139461.  
PR 18-JUN-1999. 9905-0139462.  
PR 18-JUN-1999. 9905-0139463.  
PR 18-JUN-1999. 9905-0139750.  
PR 18-JUN-1999. 9905-0139763.  
PR 21-JUN-1999. 9905-0139817.  
PR 22-JUN-1999. 9905-0139899.  
PR 23-JUN-1999. 9905-0140353.  
PR 23-JUN-1999. 9905-0140354.  
PR 24-JUN-1999. 9905-0140655.  
PR 28-JUN-1999. 9905-0140823.  
PR 29-JUN-1999. 9905-0140991.  
PR 30-JUN-1999. 9905-0141287.  
PR 01-JUL-1999. 9905-0141842.  
PR 01-JUL-1999. 9905-0142154.  
PR 02-JUL-1999. 9905-0142055.  
  
PR 06-JUL-1999. 9905-0142390.  
PR 08-JUL-1999. 9905-0142803.  
PR 09-JUL-1999. 9905-0142920.  
PR 12-JUL-1999. 9905-0142977.  
PR 13-JUL-1999. 9905-0143542.  
PR 14-JUL-1999. 9905-0143624.  
PR 15-JUL-1999. 9905-0144008.  
PR 16-JUL-1999. 9905-0144082.  
PR 16-JUL-1999. 9905-0144086.  
PR 19-JUL-1999. 9905-0144325.  
PR 19-JUL-1999. 9905-0144331.  
PR 19-JUL-1999. 9905-0144332.  
PR 19-JUL-1999. 9905-0144333.  
PR 19-JUL-1999. 9905-0144334.  
PR 19-JUL-1999. 9905-0144335.  
PR 20-JUL-1999. 9905-0144335.  
PR 20-JUL-1999. 9905-0144632.  
PR 20-JUL-1999. 9905-0144632.  
PR 21-JUL-1999. 9905-0144884.  
PR 21-JUL-1999. 9905-0144884.  
PR 21-JUL-1999. 9905-0145086.  
PR 22-JUL-1999. 9905-0145085.  
PR 22-JUL-1999. 9905-0145087.  
PR 22-JUL-1999. 9905-0145089.  
PR 22-JUL-1999. 9905-0145192.  
PR 23-JUL-1999. 9905-0145145.  
PR 23-JUL-1999. 9905-0145218.  
PR 23-JUL-1999. 9905-0145224.  
PR 26-JUL-1999. 9905-0145276.  
PR 27-JUL-1999. 9905-0145911.  
PR 27-JUL-1999. 9905-0145918.  
PR 28-JUL-1999. 9905-0145951.  
PR 02-AUG-1999. 9905-0146368.  
PR 02-AUG-1999. 9905-0146388.  
PR 03-AUG-1999. 9905-0146588.  
PR 03-AUG-1999. 9905-0147039.  
PR 04-AUG-1999. 9905-0147204.  
PR 04-AUG-1999. 9905-0147204.  
PR 05-AUG-1999. 9905-0147199.  
PR 05-AUG-1999. 9905-0147199.  
PR 06-AUG-1999. 9905-0147303.  
PR 06-AUG-1999. 9905-0147416.  
PR 09-AUG-1999. 9905-0147935.  
PR 09-AUG-1999. 9905-0147935.  
PR 10-AUG-1999. 9905-0148171.  
PR 11-AUG-1999. 9905-0148319.  
PR 12-AUG-1999. 9905-0148341.  
PR 13-AUG-1999. 9905-0148565.  
PR 13-AUG-1999. 9905-0148684.  
PR 16-AUG-1999. 9905-0149368.  
PR 17-AUG-1999. 9905-0149175.  
PR 18-AUG-1999. 9905-0149426.  
PR 20-AUG-1999. 9905-0149722.  
PR 20-AUG-1999. 9905-0149722.  
PR 20-AUG-1999. 9905-0149722.  
PR 20-AUG-1999. 9905-0149722.  
PR 23-AUG-1999. 9905-0149902.  
PR 23-AUG-1999. 9905-0149902.  
PR 23-AUG-1999. 9905-0149930.  
PR 25-AUG-1999. 9905-0150566.  
PR 26-AUG-1999. 9905-0150884.  
PR 27-AUG-1999. 9905-0151065.  
PR 27-AUG-1999. 9905-0151066.  
PR 27-AUG-1999. 9905-0151080.  
PR 30-AUG-1999. 9905-0151303.  
PR 31-AUG-1999. 9905-0151438.  
PR 01-SEP-1999. 9905-0151539.  
PR 07-SEP-1999. 9905-0151539.  
PR 10-SEP-1999. 9905-0153070.  
PR 13-SEP-1999. 9905-0153758.  
PR 15-SEP-1999. 9905-0154018.  
PR 16-SEP-1999. 9905-0154033.  
PR 20-SEP-1999. 9905-0154779.  
PR 22-SEP-1999. 9905-0155138.

PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158332.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 25-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 64.2%; Score 358.4; DB 21; Length 1493;  
 Best Local Similarity 77.8%; Pred. No. 2.7e-96;

Matches 431; Conservative 1; Mismatches 122; Indels 0; Gaps 0;

OY 1 TGAATGTCACCAATATTTATGATAGGATATATTCCTCAATTTGTCAGACGCTGC 60  
 DB 649 TGAATGTCACGCTTCTTCTATGATATCTGATATATTCCTCAATTTGTCAGATGTC 708  
 OY 61 GCCAATTTGGAATTAAGCTGCTCTATGATGATGATATTCCTCAATTTGTCAGATGTC 120  
 DB 709 GGCATTTGCGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768  
 OY 121 GCTTTATCCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 DB 769 GCTTTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828  
 OY 181 TAGAGCTATTCAGACATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 DB 829 TAGAGCTATTCAGACATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888  
 OY 241 AATTTGCAAAAGATTTTACCTGATGATTAAGATTAAGATTAAGATTAAGATTAAG 300  
 DB 889 AATTTGCAAAAGATTTTACCTGATGATTAAGATTAAGATTAAGATTAAGATTAAG 948  
 OY 301 TGAAGAAATCTGCAATGCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 360  
 DB 949 TGAAGAAATCTGCAATGCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1008  
 OY 361 CTAGGCTTCTACCTTGGAGAGCCCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 DB 1009 CTAGGCTTCTACCTTGGAGAGCCCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068  
 OY 421 CAATCTTTGGCAATGCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 480

DB 1069 CAATTTCTGGGCAACCGCTCAAGAGCTACATATATAGACAAAGGCTGCAATGACT 1128  
 OY 481 ACCACATAGGCGCTTGGGCTGATTCCTGTAATTCATCATATGCTGATTAATTC 540  
 DB 1129 TCCACATAGGAGCTTGGGCTGATTCACAGCTGACATACATGATTCGATTAATTC 1188  
 OY 541 AGTTATGCGGCGCA 554  
 DB 1189 AGTTTGGCGGCCCA 1202

RESULT 5

ID AAC36967 standard; DNA: 1335 BP.  
 AC AAC36967;  
 DT 17-OCT-2000 (first entry)  
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 15706.  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15706.  
 XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 OS Arabidopsis thaliana.  
 XX EP1033405-A2.  
 PN Arabidopsis thaliana.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PR 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 01-APR-1999; 99US-0126785.  
 PR 06-APR-1999; 99US-0127462.  
 PR 08-APR-1999; 99US-0128234.  
 PR 16-APR-1999; 99US-0128714.  
 PR 19-APR-1999; 99US-0129845.  
 PR 21-APR-1999; 99US-0130077.  
 PR 23-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134761.  
 PR 19-MAY-1999; 99US-0134768.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 21-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139455.  
PR 16-JUN-1999; 99US-0139455.  
PR 17-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0138452.  
PR 18-JUN-1999; 99US-0138452.  
PR 18-JUN-1999; 99US-0138452.  
PR 18-JUN-1999; 99US-0138460.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140354.  
PR 23-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141843.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0143055.  
PR 06-JUL-1999; 99US-0143390.  
PR 08-JUL-1999; 99US-0143803.  
PR 09-JUL-1999; 99US-0143920.  
PR 12-JUL-1999; 99US-0143977.  
PR 13-JUL-1999; 99US-0143547.  
PR 14-JUL-1999; 99US-0143622.  
PR 15-JUL-1999; 99US-0144009.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144329.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144333.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144633.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 23-JUL-1999; 99US-0145199.  
PR 23-JUL-1999; 99US-0145199.  
PR 23-JUL-1999; 99US-0145210.  
PR 26-JUL-1999; 99US-0145222.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147410.  
PR 09-AUG-1999; 99US-0147499.

PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149375.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151065.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 28-SEP-1999; 99US-0156458.  
PR 03-OCT-1999; 99US-0157177.  
PR 03-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159243.  
PR 13-OCT-1999; 99US-0159243.  
PR 13-OCT-1999; 99US-0159243.  
PR 13-OCT-1999; 99US-0159243.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 62 0%; Score 346; DB 21; Length 1335;  
Best Local Similarity 79 6%; Pred No 1,3e-97;  
Matches 409; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 45 TTGTGACGACGTGCGCAATTGCAATAACGTCTCTATTGACTGGAATTATGCC 104

[illegible]

PA	(Dupo ) DU PONT DE NEMOURS & CO E. I.	
XX		
PI	Falco SC, Fomodu LO;	
PI		
DR	WPI: 2000-182429/16.	
DR	P-PSDB: AAY44741.	
XX		
PT	Novel tetrahydrofolate metabolism enzyme used to alter the level of	
PT	tetrahydrofolate metabolism in plants and seeds	
PS	Claim 3; Pages 34-35; 37pp; English.	
XX		
CC	The present sequence is a cDNA clone wlm96.pk047.14 encoding	
CC	5,10-methylenetetrahydrofolate reductase (MTHFR). The clone was	
CC	isolated from a wlm96 cDNA library which was prepared using wheat	
CC	seedlings obtained 96 hrs. after inoculation with E. graminis.	
CC	MTHFR plays a role in the synthesis of methionine.	
CC	The present sequence is used in the	
CC	construction of a chimeric gene to alter the level of tetrahydrofolate	
CC	metabolism/enzyme in plants. The enzyme may provide target to facilitate	
CC	design and/or identification of inhibitors that may be useful as	
CC	herbicides. The polynucleotide is also useful as a source of probes for	
CC	genetically and physically mapping the genes and as markers for traits	
XX	linked to the genes.	
XX		
S0	Sequence 451 BP; 135 A; 89 C; 101 G; 117 T; 9 other;	
	Query Match 25.0%; Score 139.6; DB 21; Length 451;	
	Best Local Similarity 74.0%; Pred. No. 1.8e-31;	
	Matches 213; Conservative 0; Mismatches 70; Indels 5; Gaps 3	
QY	1 TGAATGACACCAATTTATATGATACGATATATTCCTCAAAATTTGTGAAGACTGTC 60	
Db	163 TTAATATACCCANCTTTCTATGATATACGATATCTTCTCAAGTTTGTGACGACTGC 222	
QY	61 GCCAAATTTGGATATACGTCGCTTATTTGACTCGGATTTATGCCCTTATATTTACAAAG 120	
Db	223 GTCATATTTGGTATTAACGTCGCCCTATGCTTCGCGATATATGCAATTAATTAACACAAAG 282	
QY	121 GCTTATTCGCGATGACTGGGTTTTCGAAAACAAAGATACGAGTACATTAATGAGCTG-CT 179	
Db	283 GATTTGTGCGCATGACTGATTCGCAAAACATAAAATTCACCTGATGATTTCTGCTGCT 342	
QY	180 TTAGAGCTCTTCGAAGACATGAGAAGCTGTCAA-GGCTTATGGAATTCACCTGGGAAC 238	
Db	343 TGGGTCCTCTACTTAAGACATGAGAGGATGTGAAAACATATGAGATCCACCTGTAC 402	
QY	239 TGAATTTGCAAAAAGATTTTACGTCATGATTAACATATTCGCA 283	
Db	403 TGAAGATGTTCAAAAAAATTTTGGCTGCTGATTAACATTCGCA 450	
RESULT 7		
AAC91206		
DD	AAC91206 standard; DNA; 2187 BP.	
AC	AAC91206;	
XX		
DE	20-MAR-2001 (first entry)	
XX		
DE	Human methylenetetrahydrofolate reductase gene SEQ ID NO: 1.	
XX		
KW	Human; schizophrenia; developmental disorder; spina bifida cystica;	
KW	Tourette's syndrome; bipolar illness; autism; conduct disorder;	
KW	attention deficit disorder; obsessive compulsive disorder;	
KW	chronic multiple tic syndrome; learning disorder; polymorphism; ds.	
XX		
OS	Homo sapiens.	
XX		
XX	WO200071754-A1.	
XX		
XX	30-NOV-2000.	

Mon Feb 24 06:47:47 2003

us-09-720-451-5.rng

Page 10

PF	24-MAY-2000; 2000MO-05J4J354.
XX	
PR	25-MAY-1999; 99US-03J8448.
XX	
PA	(UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX	
PI	Johnson WG, Stenroos ES;
DR	WPI; 2001-025174/03.
XX	
PT	Diagnosing a developmental disorder, e.g. schizophrenia, by forming
PT	datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)
PT	and environmental variables affecting an individual and then comparing
PT	these DS with reference DS -
XX	
PS	Disclosure; Page 114-115; 156pp; English.
XX	
CC	The present invention provides a novel method of estimating the
CC	susceptibility of an individual to a developmental disorder using genetic
CC	and environmental variables. The method can be used in the diagnosis,
CC	prevention and treatment of disorders such as schizophrenia, spina bifida
CC	cystics, Tourette's syndrome, bipolar illness, autism, conduct disorders,
CC	attention deficit disorder, obsessive compulsive disorder, chronic
CC	multiple tic syndrome and learning disorders such as dyslexia.
XX	
QQ	Sequence 2187 BP; 479 A; 655 C; 616 G; 437 T; 0 other;

	Query Match	Similarity	53.38%	Score	116.4	DB	22	Length	2187
	Best Local			Pred	No. 2	6e-24			
	Matches	293	Conservative	0	Mismatches	251	Indels	6	Gaps
QY	1	TCATTGTCACCCAAATTTATTTATGATACGAGATATATTCCTCAAAATTTGTGACAGCTGTC	60						
DB	674	TCATATCATACGACGCTTTCTTTGTGAGCGAGACACTCTTCGCTTTGTGAAGCAATGGA	733						
QY	61	GCCAAATTTGSAATACGTGCTCTTATGTACTGGAATTAATGCCATTAATTAATTAACAG	120						
DB	734	CCGAATAGGGCAATCACTTGGCCCATGTCCTCCGGGAATCTTCCCATCCAGAGCTACCACT	793						
QY	121	GCTTATTCGCAATGACTGGGTTTGTGCAAAACAAGATACCAGTCACATTAATGCTGCTT	180						
DB	794	CCCTTTCGGCAGCTGTGTAGAGCTGCACAACTGAGAGGTGCACAGAGATCAAGAGCTGA	853						
QY	181	TGAGACCTATCAGACGAATCAGAAAGCTGTAGAGCTTATGTGAATTAACCTGGAGACTG	240						
DB	854	TTGAGACCATCAAGACAGACAGATGTCGCATCCGACATATGTGCAATGCAGAGCTGGCCGTGA	913						
QY	241	AAATGTGCAAAAAGATTTTACTCATATGATGAATGAACAT---TGACATCTTAAACCTGA	297						
DB	914	GCTGTGTACAGAGCTTCTGSCCAGTGGCTTGTCGACAGCTCCACTTCTTAACACTCTTA	973						
QY	298	ATATGAGAAATCTGCATTGGCAATCTATGAACCTTGGACCTAATTGAGAGTCCAAAG	357						
DB	974	ACCGGAGAAATGGTATCACACAGAGTCTTGAACCCCTGGGGGATGTGACATGAGAGACCCA	1033						
QY	358	TTTTCAGGTCTTACCTTGGAGAGCCGCTCAAAATGTGTTTCGATTTAAAGAAAGATGTC	417						
DB	1034	GCGCTC---CCCTACCTCGGGCTCTCAGTGGCCACCACCCAAAGCGCAGAGAGATGTAC	1090						
QY	418	GTCACATCTTCTTGGCAATGAGCAAAAGATCATATCAGACAGCAATGAGTGGGAATC	477						
DB	1091	GTCACATCTTCTTGGCTCCAGACAAAGATTAATCTACGTACCCAGAGATGGAGAG	1150						
QY	478	AATTAACCAATGGGCGTTGGGGGTATTCCTGAATCCATCATATGATGTCATTAATCTGATT	537						
DB	1151	ACTTCCCTTAACGGCCGCTGGGCAATTCCTCTCCCTGCTTTGGGGAGCTGAAGACT	1210						
QY	538	ATCAGATTCAT	547						
DB	1211	ACTACTCTCTT	1220						

## RESULT 8

AA009689	standard; cDNA; 2219 bp.
ID	AA009689
AC	AA009689;
XX	
XX	15-OCT-1996 (first entry)
DE	Human methylene-tetrahydrofolate-reductase cDNA.
XX	
KW	Methylene-tetrahydrofolate-reductase; MTHFR; gene therapy;
KW	cardiovascular disease; neurological disease; folic acid metabolism;
KW	EC:1.5.1.20; enzyme; ss.
XX	
OS	Homo sapiens.
XX	
FM	Key
FT	CD8
FT	13..1980
PN	/*tag= a
XX	
XX	MO9533054-A1.
PD	07-DEC-1995.
XX	
XX	25-MAY-1995; 95WO-CA00314.
XX	
PR	26-MAY-1994; 94GB-0010620.
XX	
PA	(UYMC-) UNIV MCGILL.
XX	
PI	Goyette P, Rozen R:
XX	
DR	WPI: 1996-030565/03.
DR	P-PSDB; AAR8358.
XX	
PT	Human methylene-tetrahydrofolate reductase cDNA probe - for
PT	detection of sequence abnormalities in methylene-tetrahydrofolate
PT	reductase e.g. in cardiovascular, neurological or folic acid
PT	metabolism disorders
XX	
PS	Disclosure: Fig. 6A-6C; 66pp; English.
XX	
CC	This sequence encoding human MTHFR has been localised to chromosome
CC	1p36.3. It may be used for the construction of DNA probes which may
CC	be used for the identification of sequence abnormalities in patients
CC	with severe or mild MTHFR deficiency. The resulting probe may also
CC	be used in gene therapy to produce the MTHFR protein.
XX	
SO	Sequence 2219 bp; 501 A; 656 C; 620 G; 442 T; 0 other;

Query Match	Similarity	20.9%	Score 116.4	DB 17	Length 2219
Best Local	Similarity	53.3%	Pred. No. 2,6e-24		
Matches	293	Conservative	0	Mismatches	251
				Indels	6
				Gaps	2
QY	1	TGATGTCCACCCCAATTTATTTATGATACGGATATATTCCTCAATTTTGGACGACATGC	60		
Db	683	TCATCATCACGACGCTTTCTCTTGAGGCGACACATCTTCGCTTTGGAAAGCATGCA	742		
QY	61	GCCAAATTTGGATTAACGTGTCTATGTACCTGGAAATTAAGCCATTAAATTAACAAG	120		
Db	743	CCACACTGGGACATCTTGCCCATGTGCTCCGGATCTTTCCCATTCACGGGCTACACT	802		
QY	121	GCTTTATCCGCATGATGCGTTTGCATAAACAAGATACCAAGCTGACATTAATGGCTGCTT	180		
Db	803	CCCTTGGGAGCTTTGTAAAGCTCTCCAAAGCTGGAGGTGCCAAGGAGATCAAGGACCTGA	862		
QY	181	TAGAGCCTATCAAGACATGAAGAAGCTGTCAAGGCTTTATGSAATTCACCTGGAACCTG	240		
Db	863	TTTAGGCCAATCAAGACACAAAGATGCTGCCATCCCACTATGGCATACAGGCGCGCTGA	922		
QY	241	AAATGTGCAAAAAGATTTTATGCTCATGSAATTAAGACAT---TGCAATCTTTATACATAA	297		
Db	923	GCGTGTGCGAGGAGCTTCTGTGCCAGTGGGTGGTGCCAGGCGTCCACTTACACACCTCA	982		

• • • • •



```

FH Key Location/Qualifiers
FT mat_peptide 1..1980
XX /*tag= a
XX
XX W09533054-A1.
XX
XX 07-DEC-1995.
XX
XX 25-MAY-1995; 95WO-CA00314.
XX
XX 26-MAY-1994; 94GB-0010620.
XX
XX (UWMC-) UNIV MCGILL.
XX
XX Goyette P, Rozen R;
XX
XX WPI: 1996-030565/03.
XX
XX P-PSDB: AAT09694.
XX
XX Human methylene:tetra:hydro:folate reductase cDNA probe - for
XX detection of sequence abnormalities in methylene:tetra:hydro:folate
XX reductase e.g. in cardiovascular, neurological or follic acid
XX metabolism disorders
XX
XX PS Disclosure: Fig.1A-1F; 66pp; English.
XX
XX CC This sequence encoding human MTHFR has been localised to chromosome
XX 1p36.3. It may be used for the construction of DNA probes which may
XX be used for the identification of sequence abnormalities in patients
XX with severe or mild MTHFR deficiency. The resulting probe may also
XX be used in gene therapy to produce the MTHFR protein.
XX
XX SQ Sequence 2220 BP; 501 A; 657 C; 620 G; 442 T; 0 other;

Query Match 20.9%; Score 116.4; DB 17; Length 2220;
Best Local Similarity 53.3%; Pred No 2,6e-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

QY 1 TGATTTCACCGCAATTTATATGATATATTCCTCAAAATTGTGACGACTGC 60
DB 663 TCATCATACGACGAGCTTCTTGTGAGCTACACATCTCCGCTTGTATGAGCATGCA 742
QY 61 GCCAAATTGGATTAAGTGTCTTGTGATGGAATTTATGTCCTTATATTTACAAAG 120
DB 743 CCGACATGGCATCACTTGTCCCATGCTCCCGGATCTTCCATCAGGCTACCACT 802
QY 121 GCTTATCGCATGACTGGGTTTGCACAAACAAAGATACAGCTGACATTATGCTGCTT 180
DB 803 CCTTGGCAGCTTGTGAAGCTGTCCAAAGCTGAGAGTGCACAGAGATAMAGACGTGA 862
QY 181 TAGAGCCTATCAAGACATGAMAGACCTGTCAGGCTTATGAAATTCACCTGGGACTG 240
DB 863 TTGAGCCCAATCAAGACATGCTGTCATCCGACATATGAGTGCATGAGTGCCTGTA 922
QY 241 AATATGTCAMAAAGATTTTACCTGATGATTAAGACAT--TGCATCTTATACCTPA 297
DB 923 GCCTGTCCAGAGCTGTGGCAGTGGCTGTGGTGCAGGCTCCACTTCTACACCTCA 982
QY 298 ATATGAGAAATCTGATTCATTAATAATGAACTGAGCTTATGAGAGTGCCTCAAG 357
DB 983 ACCGCGAGATGGCTACACACAGAGTGTGAAGGCTGGGATGTGACATGAG--GATC 1039
QY 358 TTTCTAGGTCCTTACCTTGAGAGAGCGCTGCAAAATGTTTCCGTTAAAGAAAGATGCC 417
DB 1040 CCAGGCGTCCCTACCTCGGCTCTCAGTCCGCCAACCCCAACGCCGAGAGAGAAATGTAC 1099
QY 418 GTCCAACTTTTGGGCAATGACGACAAAGAAAGCTACATATCAAGACCATAGATGGATC 477
DB 1100 GTCCCACTTTTGGGCTCCAGACCAAGATTTACCTTACCTACCCAGAGATGGAGCG 1159
QY 478 AATACCAATGGGGTGGGTATTCCTGTAATCATCATATGAGCTTATATCTGATT 537
DB 1160 AGTTCCTTAACGGCCGCTGGGCAATTCCTTCCCTGCTTGGGAGCTGAAGACT 1219

```

```

QY 538 ATCACTTCAT 547
DB 1220 ACTACCTCTT 1229

RESULT 11
ID AAAS0633 standard; cDNA: 2220 BP.
AC AAAS0633;
XX
XX 19-DEC-2000 (first entry)
XX
XX Human methylenetetrahydrofolate reductase cDNA.
XX
XX Methyleneetetrahydrofolate reductase; MTHFR; human; follic acid;
XX cardiovascular disorder; cancer; neuroblastoma;
XX colorectal carcinoma; osteoporosis; neural tube
XX neurological disorder; gene therapy; diagnosis; chromosome 1p36.3;
XX ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT CDS 13..1983
XX FT /*tag= a
XX FT /*tag= "1.5.1.20"
XX FT replace(167,A)
XX FT /*tag= b
XX FT replace(482,A)
XX FT /*tag= c
XX FT replace(559,T)
XX FT /*tag= d
XX FT replace(677,T)
XX FT /*tag= e
XX FT replace(692,T)
XX FT /*tag= f
XX FT replace(764,T)
XX FT /*tag= g
XX FT replace(792+1,A)
XX FT /*tag= h
XX FT replace(985,T)
XX FT /*tag= i
XX FT replace(1015,T)
XX FT /*tag= j
XX FT replace(1018,T)
XX FT /*tag= k
XX FT replace(1298,C)
XX FT /*tag= l
XX FT replace(1317,C)
XX FT /*tag= m
XX
XX PN W0200052205-A2.
XX
XX PD 08-SEP-2000.
XX
XX PF 28-FEB-2000; 2000WO-1B00442.
XX
XX PR 01-MAR-1999; 99US-0258928.
XX
XX PA (UWMC-) UNIV MCGILL.
XX
XX PI Rozen R, Goyette P;
XX
XX DR WPI: 2000-572192/53.
XX
XX F-PSDB: AAT96186.
XX
XX CDNA probe for the human methylenetetrahydrofolate reductase (MTHFR),
XX useful in gene therapy and for diagnosing or treating MTHFR deficiency
XX which is associated with cardiovascular disorders or cancer
XX
XX PS Claim 1; Fig 1A-F; 93pp; English.

```



QY	61	GCAGAAATGGGATTAACGGTCTCTTTGACACGGAAATTAAGCCATAATTAATTAACAG	120
Db	743	CCGACATGGGCAATCAATTGGCCATGGTCCCGGGATTTTCCATCCAGGGCTACACCT	802
QY	121	GGTTATTCGGCATGACACTGGGTTTGGCAAAACAATATCCAGCTGACATTAATGGCTCTT	180
Db	803	CCCTTCGGCAGCTTGTGAAGCTGTCCACAGCTGGAGGTGCCACAGAGATCAAGACGTGA	862
QY	181	TGAGACCTATTCAGGACATATAAGACTGTCAAGGCTTATGGAATTCACCTCGGGAAGT	240
Db	863	TTGAGCCATCTAAAGACCAACATCTGTGCCATCCGACATATGACATCGAGCTGGCCGTGA	922
QY	241	AAATGTCAAAAAGATTTTATGCTCATGGAATTAAGACAT---TGACCTTTATACCTAA	297
Db	923	GCCTGTCCAGGAGCTCTGACACAGTGGCTGTGTCCAGGCTCTCCATCTTACACCTCA	982
QY	298	ATATGAGAAATCTGCATTGGCAATTAATAAGCACTTGGCTCAATTAAGAGATCCAAAG	357
Db	983	ACCGGAGATGGGTACACAGATGGTGTAAACGGCTGGGAGATGTGACATAGAGACCCA	1042
QY	358	TTCTAGGTCCTTACCTTGGAGAGCGCCCTGCCAATGTTCCTCGTTAAAGAAATGTCC	417
Db	1043	GGCGTC---CCCTACCCCTGGGCTCTCAGTGGCCACCCCAAGCCGACGAGGAAGATGTAC	1099
QY	418	GTCACATCTTTTGGGCAATGCAGCAAAAAGCTACATATCAAGAGACATAGAGTGGGATC	477
Db	1100	GTCGCATCTTCTGGGGCTGTCCAGACCAAAAGTTATATCTCCGTACCCAGAGGGGAGG	1155
QY	478	AATTAACCAACATGGCGTGTGGGGTATTCCTGTATATCCATATATATGGTGCATTATCTATT	537
Db	1160	ACCTTCCCTTAACGGCCGCGTGGGCAATTCCTCTTCCCTGGCTTGGGGACCTGAAAGTACT	1215
QY	538	ATCAGTTCAT	547
Db	1220	ACTACCTCTT	1229

Result	Accession	Standard	CDNA	BP
13	ABK13501	ABK13501	standard	2220
XX	XX	ABK13501		
AC	XX	ABK13501		
AD	XX			
AE	XX			
AF	XX			
AG	XX			
AH	XX			
AI	XX			
AJ	XX			
AK	XX			
AL	XX			
AM	XX			
AN	XX			
AO	XX			
AP	XX			
AQ	XX			
AR	XX			
AS	XX			
AT	XX			
AX	XX			
AY	XX			
AZ	XX			
BA	XX			
BB	XX			
BC	XX			
BD	XX			
BE	XX			
BF	XX			
BG	XX			
BH	XX			
BI	XX			
BJ	XX			
BK	XX			
BL	XX			
BM	XX			
BN	XX			
BO	XX			
BP	XX			
BQ	XX			
BR	XX			
BS	XX			
BT	XX			
BU	XX			
BV	XX			
BW	XX			
BX	XX			
BY	XX			
BZ	XX			
CA	XX			
CB	XX			
CC	XX			
CD	XX			
CE	XX			
CF	XX			
CG	XX			
CH	XX			
CI	XX			
CJ	XX			
CK	XX			
CL	XX			
CM	XX			
CN	XX			
CO	XX			
CP	XX			
CQ	XX			
CR	XX			
CS	XX			
CT	XX			
CU	XX			
CV	XX			
CW	XX			
CX	XX			
CY	XX			
CZ	XX			
DA	XX			
DB	XX			
DC	XX			
DD	XX			
DE	XX			
DF	XX			
DG	XX			
DH	XX			
DI	XX			
DJ	XX			
DK	XX			
DL	XX			
DM	XX			
DN	XX			
DO	XX			
DP	XX			
DQ	XX			
DR	XX			
DS	XX			
DT	XX			
DU	XX			
DV	XX			
DW	XX			
DX	XX			
DY	XX			
DZ	XX			
EA	XX			
EB	XX			
EC	XX			
ED	XX			
EE	XX			
EF	XX			
EG	XX			
EH	XX			
EI	XX			
EJ	XX			

```

PR      12-JUN-2000; 2000US-0692595.
PA      (UYMC-) UNIV MCGILL.
XX
PI      Rozen R;
DR
XX      WPI; 2002-130741/17.
XX      P-PDBJ; AA075413.
PT      Diagnosing subjects at risk for or suffering from a psychosis,
PT      particularly schizophrenia comprises determining the presence of a
PT      heterozygous methylenetetrahydrofolate reductase mutant allele in the
PS      disclosure; Fig 1A-F; 102pp; English.
XX
XX      The invention describes a method of diagnosing a psychosis in a subject,
CC      or a risk for or propensity to psychosis in a subject comprising
CC      determining the presence of a heterozygous methylenetetrahydrofolate
CC      reductase (MTHFR) mutant allele in a nucleic acid sample obtained from a
CC      subject. The method is useful for diagnosing subjects at risk of, or
CC      suffering from a psychosis, particularly schizophrenia but includes
CC      manic-depressive disease, organic psychotic disorders, postpartum
CC      psychosis, senile psychosis, transference psychosis and acute idiopathic
CC      psychotic illnesses. A new pharmacological composition is used for
CC      treating amyotrophic lateral sclerosis, anxiety, dementia depression,
CC      epilepsy, Huntington's disease, migraine, schizophrenal diseases, multiple
CC      sclerosis, pain, Parkinson's disease, schizoid personality disorders, stroke. Deficiency may be associated with diseases including pharynges, or
CC      cardiovascular disorders, cancer (e.g. neuroblastoma and colorectal
CC      carcinoma), osteoporosis, metabolic or endocrine disease, inborn errors
CC      of metabolism, inflammation, immune disorders, neoplastic disease and
CC      renal disease. This sequence encodes human methylenetetrahydrofolate
CC      reductase (MTHFR) which maps to chromosome 1p36 (EC number 1.5.1.20),
XX      described in the method of the invention.
XX
SO      Sequence 2220 BP; 501 A; 658 C; 620 G; 441 T; 0 other;

Query Match                20.9%; Score 116.4; DB 24; Length 2220;
Best Local Similarity    53.3%; Pred. No. 2.Ee-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps .2.

OY      1 TGATTTCACCAATATTATTATGATGCAGATATATCTTCAAAATTTGTGAACGACTGTC 60
           ||||| |   ||||| |   ||||| |   ||||| |   ||||| |   ||||| |
Db       TCATCATCACCAGAGTTTTCTTGAGGGTGACACACTCTTCCGCTTGTGAAGGATGCA 742

OY      61 GCACAATTGAGATAACGTCGTCATTGTACCGTAGTAATTATGCCATTATATTATCAAGG 120
           ||||| |   ||||| |   ||||| |   ||||| |   ||||| |   ||||| |
Db       CCAGACATGGGCATCATCTTGCCCCATCGTCCC GGAGATCTTCCCATTCAGGGCTCACT 802

OY      121 GCTTTTCCGCATGATCGGGGTTTTTGAAAACAAAGATACCAAGCTGCATTTAATGGCTGTT 180
           ||||| |   ||||| |   ||||| |   ||||| |   ||||| |   ||||| |
Db       CCTTCGGCAGCTGTGTGAAGCTGTCCAAGCTGAGAGTGCCACAGSGAGATCAAGAACGTGA 862

OY      181 TTAGAAGCTPACAGAGCAATGAGAACGTCGTACAGCGTTATGGAATGATCACCGGSAAGTG 240
           ||||| |   ||||| |   ||||| |   ||||| |   ||||| |   ||||| |
Db       TTAGAGCATATCAAAGACAGATGCTGCACATGCCAATCGCAATGACATCAGAGCGCCGTGA 922

OY      241 AATATGCAAAAAAGATTTTACGTCATGGAATTAAGACAT---TGCATCTTTATACCTAA 297
           ||||| |   ||||| |   ||||| |   ||||| |   ||||| |   ||||| |
Db       GCGCTGGACAGAGGCTCTTGCCAGAGGCTGTGGTGTGCCAGGCGCTCCATCTTACACCTCA 982

OY      298 ATATGAGAAATCTGCATGTGCATACTAATGAACCTTGGSCCTTAATGAAGATCCAAAG 357
           ||||| |   ||||| |   ||||| |   ||||| |   ||||| |   ||||| |
Db       ACCGCAAGATGGCTACACAGAGGTGCTGAACCGCTGGGGAGATGTGAGTGAAGGAGCCCA 1042

OY      358 TTTTCTAGGCTCTTACTTTGGAGAGCGCCCTCAAAATGTTTTCCGTGTTAAAGAGATGTC 417
           ||||| |   ||||| |   ||||| |   ||||| |   ||||| |   ||||| |
Db       GAGGCTC---CCCTTACCCCTGGGCTCTTACGTGCGCCACCCCAAGGCCCCAGAGAGATGTAC 1043

OY      418 GTCCATCTTTTGGGCAAAATGAGCAAAAAGCTGACATATCAAGSACATAGAGATGGATC 477
           ||||| |   ||||| |   ||||| |   ||||| |   ||||| |   ||||| |

```

Db 1100 GTCCATCTTGTGGGCTCCAGACCAAGAGTACATCTACCGTACCCAGAGATGGGACG 1159  
Qy 478 AATACCCACATGGGGCTGTGATTCCTGATTCATCATCATATGTCATTATTCGATT 537  
1160 AGTTCCTAACGGCGCTGGGCAATTCCTTCCCTGCTTGGGAGCTGAAGGACT 1219  
Qy 538 ATCAGTTCAT 547  
Db 1220 ACTACCTCTT 1229

RESULT 14  
ABK13539  
ID ABK13539 standard; cDNA; 2220 BP.  
XX ABK13539;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human methyltetrahydrofolate reductase (MTHFR) G167A allele.  
XX  
KW Methyltetrahydrofolate reductase; MTHFR; neuroleptic;  
KW neuroprotective; tranquilizer; nootropic; antidepressant;  
KW anticonvulsant; cerebroprotective; amyotrophic lateral sclerosis;  
KW anxiety; dementia; depression; epilepsy; Huntington's disease;  
KW migraine; demyelinating disease; multiple sclerosis; pain;  
KW Parkinson's disease; psychosis; stroke; cardiovascular disorder;  
KW cancer; osteoporosis; metabolic disease; endocrine disease;  
KW inborn error of metabolism; inflammation; immune disorder; human;  
KW neuroblastoma; colorectal carcinoma; neoplastic disease; renal disease;  
KW ss; EC number 1.5.1.20; mutant.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1983  
FT /tag= a  
FT /product= "MTHFR"  
FT /note= "Methyltetrahydrofolate reductase"  
FT replace(167,g)  
FT /tag= b  
XX  
PN W0200196598-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 12-JUN-2001; 2001WO-CA00867.  
XX  
PR 12-JUN-2000; 2000US-0592595.  
XX  
PA (UYMC-) UNIV MGSILL.  
XX  
PI Rozen R;  
XX  
DR WPI; 2002-130741/17.  
DR P-PSDB; AAU5421.  
XX  
PT Diagnosing subjects at risk for or suffering from a psychosis,  
PT particularly schizophrenia comprises determining the presence of a  
PT heterozygous methyltetrahydrofolate reductase mutant allele in the  
PT subject  
XX  
PS Claim 4; Page -; 102pp; English.  
XX  
CC The invention describes a method of diagnosing a psychosis in a subject,  
CC or a risk for or propensity to psychosis in a subject comprising  
CC determining the presence of a heterozygous methyltetrahydrofolate  
CC reductase (MTHFR) mutant allele in a nucleic acid sample obtained from a  
CC subject. The method is useful for diagnosing subjects at risk of, or  
CC suffering from a psychosis, particularly schizophrenia but includes  
CC manic-depressive disease, organic psychotic disorders, psychosis in  
CC alcohol or drug intoxication, postinfection psychosis, postpartum

CC psychosis, senile psychosis, traumatic psychosis and acute idiopathic  
CC psychotic illnesses. A new pharmaceutical composition is used for  
CC treating amyotrophic lateral sclerosis, anxiety, dementia, depression,  
CC epilepsy, Huntington's disease, migraine, demyelinating disease, multiple  
CC sclerosis, pain, Parkinson's disease, schizophrenia, psychoses, or  
CC stroke. Deficiency may be associated with diseases including  
CC cardiovascular disorders, cancer (e.g. neuroblastoma and colorectal  
CC carcinoma), osteoporosis, metabolic or endocrine disease, inborn errors  
CC of metabolism, inflammation, immune disorders, neoplastic disease and  
CC renal disease. This sequence encodes a mutant human  
CC methyltetrahydrofolate reductase (MTHFR, EC number 1.5.1.20), in which  
CC nucleotide 167 has been altered to produce a mutant allele, described in  
CC the method of the invention.  
CC Note: This sequence does not appear in the specification but has been  
CC created from the wild type sequence (ABK13501) using information given  
CC in claim 4 of the invention.  
XX  
SQ Sequence 2220 BP; 502 A; 658 C; 619 G; 441 T; 0 other:  
Query Match 20.9%; Score 116.4; DB 24; Length 2220;  
Best Local Similarity 53.3%; Pred. No. 2,6e-24;  
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;  
Qy 1 TGATGTCACCCAAATATTATGATGATGATATATTCCTCAAAATTGGAAGCACTGC 60  
Db 683 TCATCATCAGCAGGCTTTCTTGAAGGCTGACACATCTTCGCTTGAAGGACATGCA 742  
Qy 61 GCCAAATTGGAATACGTCGTCTATTTGACCTGGAAATATGCCCATTAATTAACAAG 120  
Db 743 CCAGCAATGAGCAATCCTGCCCCCAATGCTCCCGGAGATCTTCCCAATCCAGGGCTACCT 802  
Qy 121 GCTTTATCCGCAATGACTGGGTTTGCAAAACAATACAGCTACATTAATGCTGCTT 180  
Db 803 CCTTCGGCAGCTGTGAAGCTGTCCAAAGCTGAGGTCACAGAGATCAAGGACGTGA 862  
Qy 181 TAAAGCCTATCAAGACATGAAGAAGCTGCAAGCTTAATGAATCACTGGGAACTG 240  
Db 863 TTGAGCCAAATCAAGACATGATGTGCATCCGCAACTATGACATGACGTGCGCTGA 922  
Qy 241 AAATGTGCAAAAAGATTATTAAGTCAATGAATTAAGACAT--TGACATTTATACATAA 297  
Db 923 GCCGTGCGAGAGCTTCGCGCAGTGGCTGTCGAGGCTCCACTCTACACCTCA 982  
Qy 298 ATATGAGAAATGATGATTCGCAATACATGAACCTTGCGCTAATGAAGTCCAAAG 357  
Db 983 ACCGCGAATGCTTACACACAGAGTGTGAAGCCCTGGGGATGTGACTGAGACCCCA 1042  
Qy 358 TTTCATAGTCTTACCTTGAGAGCGCCCTGCAAAATGTTTCGCTTAAGAAGATGTC 417  
Db 1043 GGGCTC--CCCTACCTGGGCTCTCAATGCTCCACCCCAAGSCCGAGAGATGTATAC 1099  
Qy 418 GTCCATCTTTTGGGCAATGACCAAAAAGCTATATCAAGACCATATGATGGATC 477  
Db 1100 GTCCATCTTTTGGGCTGCTCAAGACCAAAAGTTTATCATCTACCTACCCAGATGGAGCG 1159  
Qy 478 AATACCCACATGGGCTGTGGGCTGATTCCTGTAATCATATATATGTCATTTATGTT 537  
Db 1160 AGTTCCTAACGGCGCTGGGCAATTCCTTCCCTGCTTGGGAGCTGAAGGACT 1219  
Qy 538 ATCAGTTCAT 547  
Db 1220 ACTACCTCTT 1229

RESULT 15  
ABK13540  
ID ABK13540 standard; cDNA; 2220 BP.  
XX ABK13540;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human methyltetrahydrofolate reductase (MTHFR) G482A allele.

XX Methylentetrahydrofolate reductase; MTHFR, neuroleptic;  
KW neuroprotective; tranquiliser; nootropic; antidepressant;  
KW anticonvulsant; cerebroprotective; amyotrophic lateral sclerosis;  
KW anxiety; dementia; depression; epilepsy; Huntington's disease;  
KW migraine; demyelinating disease; multiple sclerosis; pain;  
KW Parkinson's disease; psychosis; stroke; cardiovascular disorder;  
KW cancer; osteoporosis; metabolic disease; endocrine disease;  
KW inborn error of metabolism; inflammation; immune disorder; human;  
KW neuroblastoma; colorectal carcinoma; neoplastic disease; renal disease;  
KW ss; EC number 1.5.1.20; mutant.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FH CDS 1..1983  
FT /\*tag= a  
FT /\*product= "MTHFR"  
FT /note= "Methylentetrahydrofolate reductase"  
FT mutation replace(482,G)  
FT /\*tag= b  
XX  
PN W0200196598-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 12-JUN-2001; 2001WO-CA00867.  
XX  
PR 12-JUN-2000; 2000US-0592595.  
XX  
PA (UYMC-) UNIV MCGILL.  
XX  
PI Rozen R;  
XX  
DR WPI: 2002-130741/17.  
DR P-PSDB; AAU75422.  
XX  
PT Diagnosing subjects at risk for or suffering from a psychosis',  
PT particularly schizophrenic comprises determining the presence of a  
PT heterozygous methylentetrahydrofolate reductase mutant allele in the  
PT subject -  
XX  
XX  
PS Claim 4; Page -; 102pp; English.  
XX  
CC The invention describes a method of diagnosing a psychosis in a subject,  
CC or a risk for or propensity to psychosis in a subject comprising  
CC determining the presence of a heterozygous methylentetrahydrofolate  
CC reductase (MTHFR) mutant allele in a nucleic acid sample obtained from a  
CC subject. The method is useful for diagnosing subjects at risk of, or  
CC suffering from a psychosis, particularly schizophrenia but includes  
CC manic-depressive disease, organic psychotic disorders, psychosis in  
CC alcohol or drug intoxication, postinfection psychosis, postpartum  
CC psychosis, senile psychosis, traumatic psychosis and acute idiopathic  
CC psychotic illnesses. A new pharmaceutical composition is used for  
CC treating amyotrophic lateral sclerosis, anxiety, dementia, depression,  
CC epilepsy, Huntington's disease, migraine, demyelinating disease, multiple  
CC sclerosis, pain, Parkinson's disease, schizophrenia, psychoses, or  
CC stroke. Deficiency may be associated with diseases including  
CC cardiovascular disorders, cancer (e.g. neuroblastoma and colorectal  
CC carcinoma), osteoporosis, metabolic or endocrine disease, inborn errors  
CC of metabolism, inflammation, immune disorders, neoplastic disease and  
CC renal disease. This sequence encodes a mutant human  
CC methylentetrahydrofolate reductase (MTHFR, EC number 1.5.1.20), in which  
CC nucleotide 482 has been altered to produce a mutant allele, described in  
CC the method of the invention.  
CC Note: This sequence does not appear in the specification but has been  
CC created from the wild type sequence (ABR13501) using information given  
CC in claim 4 of the invention.  
XX  
XX Sequence 2220 BP; 502 A; 658 C; 619 G; 441 T; 0 other;

Best Local Similarity -53.3%; Pred. NO. 2.6e-24;  
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

OY	1	TGATTTGACCCCAATTAATTTATGATACAGCAATATATCTCAATTTATGTGAACAGCTGTC	60
Db	683	TGATTCATCAGCAGAGCTTTCTTTTGAGGCTGACACATCTCTCCGTTTGTGAAGGCATGCA	742
OY	61	GCCAAATGTGAAATACGTCGTCATTGTACGTCGAAATTAATCCCATTAATTAATTAACAG	120
Db	743	CGCACATGTGSCATCAGCTTGGCCCATGCTGCCGGGATCTTTCCATGCAAGGCTACACT	802
OY	121	GCTTTATCCGATGACTGAGTGGTTTGCAAAACAAAGATACCACTGACATTATGCTGCTT	180
Db	803	CCCTTTGGGAGCTGTGTGAACCTGTCCAAACCTGGAGAGTCCCAAGAGATCAAGAGAGTGA	862
OY	181	TAGAGCCATTCAGAGACAAATGAAAGAGCTGTCAAGGCTTAATGSAATTAACCTGGAGCTG	240
Db	863	TTGAGCCATTCAAAGACAAAGATGTGTCATCCGACATATGATGATGAGACTGCCGCTGA	922
OY	241	AAATGTGCAAAAAGTTTTGTAGCTATGSAATTAACAGAT---TGCACTCTTAACACTAA	297
Db	923	GCGTGTGCCAAGSAGCTTCTGCGCAGTGGCTTGCTGGCCAGGCGCTCACACTTCTACACCTCA	982
OY	298	ATATGGAAGAAATCTGCAATGTGSCAAATACTAATGACCTTGGCTTAATGGAAGATCCAAAG	357
Db	983	ACCGCAGATGTGCTACACAGAGGTGCTAAGAGCGCTGGAGTGTGACTGAGAGACCCA	1042
OY	358	TTTTAGTGCTTAATCTTGGAGACGCCCTTGCAAAATGTTTTCCGTGTTAAAGAGATGTCC	417
Db	1043	GCGGCT---CCCTACCCCTGGGCTCTTAAGTGCCCAACCCMAAGCGCCAGAGAGAGATGTC	1099
OY	418	GTCGCATCTTTTTGGGCAAAATGCAACCAAAAGGTACATATCAAGAGCCATNGATGGGATC	477
Db	1100	GTCGCATCTTTTGGGCTCCGACACCAAAAGTTACATCTACCGTACCCAGAGATGGGAGC	1159
OY	478	AATACCCAAATGGGGTGTGGGGATGATCCGTAAATCAATCATPATGTGCTATTATCGAAT	537
Db	1160	AGTTCCCTCAAGCGCGCTGGGGCAATCTCCTTCGCCCTTTGGGAGAGCTGAAGAGCT	1219
OY	538	ATCAGTTTCAT	547
Db	1220	ACTGACTCTT	1229

Search completed: February 12, 2003, 10:32:27  
Job time : 288 secs

Query Match	20.98;	Score 116.4;	DB 24;	Length 2220;
-------------	--------	--------------	--------	--------------



Qy	358	TTTTCAGGTCCTTACCTTGAGAGCGCCCTGCAAATGTTTTCCGTGTTAAAGAAATATCC	417
Db	1031	GGAGC---CCCAACCCCTGGCTCTCAGTGCCCAACCCAGAGCCGAGAGGAGATGTAC	1087
Qy	418	GTCCATATCTTTTGGGCAATTCGACAAAAAGTATATTCMAAGACATAGATGGAGATC	477
Db	1088	GTCCATCTCTTCTGGGCTCCGACACAAAGATTATATCTACCTACCCGAGAGTGGAGAC	1144
Qy	478	AATACCAATGGGGTGGGGTATCTCTGTAAATCATCATATGGTGATTAATCATAT	537
Db	1148	AGTTTCCCTAAGCGCCGCTGGGACATATCTCTCCCTGCGCTTGGGAGCTGAAGGACT	1207
Qy	538	ATCAGTTCAT	547
Db	1208	ACTACCTCTT	1217

```

RESULT 2
US-09-318-448-1
? Sequence 1, Application US/09318448
? Patent No. 6210950
? GENERAL INFORMATION:
? APPLICANT: Johnson, William G.
? APPLICANT: Steenros, Edward S.
? TITLE OF INVENTION: METHODS FOR DIAGNOSING,
? TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
? FILE REFERENCE: 601-1-057
? CURRENT APPLICATION NUMBER: US/09/318,448
? CURRENT FILING DATE: 1999-05-25
? NUMBER OF SEQ ID NOS: 46
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 1
? LENGTH: 2187
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-318-448-1

```

Query Match	20.98;	Score 116.4;	DB 4;	Length 2187;
Best Local Similarity	53.38;	Pred. No. 2.1e-27;		
Matches 293; Conservative	0;	Mismatches 251;	Indels 6;	Gaps 2;

Qy	1	TCATGTGACC	CAATTTATTTATGATACG	GAATTTTTC	CAATTTTGG	AGAC	GTG	60
Db	674	TCATCATC	ACAGAGCTTTCTT	TGAGCTG	ACACATCTTCCCTTT	TGGAAG	CAATGCA	733
Qy	61	GCCAAAT	TGGAAATACG	TGCTCTTTGT	ACTCG	GAATTTAG	CCCATTAATTTAC	120
Db	734	CCGAC	ATGGGCATACCTTG	CCCCCATG	CTCCCGGGATCTTTTCC	CAATCCAG	GGCTAC	793
Qy	121	GCTTAT	CCGATGACTGG	TTTTGG	AAACAAAGATAC	CGTACATTTAG	CTGCTT	180
Db	794	CCCTTC	CGCAGCTGTG	AGCTGTCC	CAACCTG	AGAGTCC	CAAGAGATCA	853
Qy	181	TAGAG	CCCTATCA	AGCAATATAG	AGCTGTGA	AGGCTTTG	GAATATCA	240
Db	854	TTGAG	CCCATCAAG	CAAGCATGTG	CGCATCCG	CACTACG	AGCTG	913
Qy	241	AAATG	CGAAAAGATTTT	TACTCAT	TGGAATTTAG	ACAT---TGC	ATCTTATCA	297
Db	914	GCTGTG	CTCAGAGGCTCTG	GCACAGTGTG	TGTGTG	TGAGGCTTCC	CACTTTCAC	973
Qy	298	ATATG	GAGAAATTCAT	TGCAATG	CAATCTT	GGCTTAATTA	AGTCC	357
Db	974	ACCCG	GAGATGGCTAC	ACACAGGTCT	GAACCGCTTG	GGATGTG	ATGAGAC	1033
Qy	358	TTTTC	AGTCCCTTAC	CTTGG	AGACGGCTCG	CAAAATGTTTTCC	TGTAA	417
Db	1034	GGCGT---	CCCTAC	CCCTGG	CTCTCAGT	GCCAC	CCCAAGCCCG	1090
Qy	418	GTC	CAATCTTTTGG	CAATG	ACCAAAAAG	CTACATAT	TAG	477
Db	1091	GTC	CAATCTTTTGG	CAATG	ACCAAAAAG	CTACATAT	TAG	1150

Oy	478	AATACCACCACTGGGCGGTGATTCCGTGAATCATCATGTGTCATTTATGCAGT	537
Dd	1151	AATTGCCCAAGGCGCCTGGGGCAATTCTCTCTCCCTGGCTTGGGAAAGTGAAAGACT	1210
Oy	538	ATCATCTTCAT	547
Dd	1211	ACTACCTCTT	1220

US-08-738-000-3

RESULT 3

Sequence 3, Application US/08738000

Patent No. 6074821

GENERAL INFORMATION:

TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE

TITLE OF INVENTION: REDUCTASE

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER & JACKSON

STREET: Continental Plaza - 411 Hackensack Avenue

City: Hackensack

STATE: New Jersey

COUNTRY: U.S.A.

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/738, 000

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: NO PCT/CA95/00314

FILING DATE: 25-MAY-1995

PRIOR APPLICATION DATA: GB 9410620.0

APPLICATION NUMBER: GB 9410620.0

FILING DATE: 26-MAY-1994

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 2219 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 13..1983

US-08-738-000-3

Query Match	20.98;	Score 116.4;	DB 3;	Length 2219;
Best Local Similarity	53.38;	Pred. No. 2.1e-27;		
Matches 293; Conservative	0;	Mismatches 251;	Indels 6;	Gaps 2;

[illegible]

```

OY 241 AATGTGCAAAAAGATTTTACCTCATGGAATTAAGACAT---TGCATCTTTATACACTAA 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 923 GCCTGTGCCAGAGAGCTTCTGTGCCAGTGTGTGCCAGAGGCTCTACCTTTCTACACCTTCA 982
OY 298 ATATGAGAAATCTGCATTGGCAATATCATATGACCTTGGCTTAATGGAAGTCCAAAG 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 983 ACCGCGAGATGGCTACACAGAGGTGTGAAGGGCTGGAGATGTGACTGAGAGGCCCA 1042
OY 358 TTTCTAGTCTTACCTTGTGAGACGCCCTGCACAAATGTTTCCGTGTTAAAGAATGTCC 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1043 GGGCTC---CCCTACCCCTGGGCTCTCAGTGCCTCCACCCCAAGCGCCGAGAGAGATGTAC 1099
OY 418 GTCCAACTTTTGGGCAAAATCGACAAAAGCTACATATCAAGAGCAATAGATGGGATC 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1100 GTCCCACTTTCTGGGCTCTCCAGACCAAGATTAATCTACCGTACCCAGAGATGGGACG 1159
OY 478 AATACCAATGCGGCTTGGGGGTGATTCCTGTATCATCATATGATGTGATTCGATT 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1160 AGTTCCCTTAAGCGCGCTGGGCAATTCCTCTTCCCTGGCTTTGGGAGCTGAAGACT 1219
OY 538 ATCAGTTTCAT 547
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1220 ACTACCTCTT 1229

```

RESULT 4

```

US-09-720-928-3
; Sequence 3, Application US/09258928
; Patent No. 6218120
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/005002
; CURRENT APPLICATION NUMBER: US/09/258,928
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/738,000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9410620.0
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (113)...(1983)
US-09-720-928-3

```

Query Match 20.9%; Score 116.4; DB 4; Length 2219;  
 Best Local Similarity 53.3%; Pred. No. 2.1e-27;  
 Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

```

OY 1 TGATGTACCAATTTATATGATAGCATATATTCCTCAAAATTTGTAGACACTGTC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 663 TCATCATCACGACGCTTTCTTGTAGGCTGTACACATCTTCCGCTTGTGAAGCATGCA 742
OY 61 GCCAAATTTGGAATAGGTGTCTTGTACCTGGAATTTATGCCATTAATTAATTAACAAG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 743 CCGACATGGGCAATCTGCCCCCATCGTCCCGGAGTCTTCCATCCAGGGCTACCACT 802
OY 121 GCTTATCCGATGACTGGGTTTGTGCAAAACAAAGATACAGCTGACATTAATGCTGCTT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 803 CCCTTGGGAGCTTGTGAAGCTGTCCAAAGCTGAGAGTGCACAGAGATCAAGAGAGTGA 862
OY 181 TAGAGGCTATCAGAGCATGAGAGAGCTGTCAAGGCTTAATGGAATTCACCTGGGAATG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 863 TTAGGCCAATCAAGACAAACATGTGCTGCTCGCAACTATGATGATGAGCTGGCGGTGA 922

```

```

OY 241 AATGTGCAAAAAGATTTTACCTCATGGAATTAAGACAT---TGCATCTTTATACACTAA 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 923 GCCTGTGCCAGAGAGCTTCTGTGCCAGTGTGTGCCAGAGGCTCTACCTTTCTACACCTTCA 982
OY 298 ATATGAGAAATCTGCATTGGCAATATCATATGACCTTGGCTTAATGGAAGTCCAAAG 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 983 ACCGCGAGATGGCTACACAGAGGTGTGAAGGGCTGGAGATGTGACTGAGAGGCCCA 1042
OY 358 TTTCTAGTCTTACCTTGTGAGACGCCCTGCACAAATGTTTCCGTGTTAAAGAATGTCC 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1043 GGGCTC---CCCTACCCCTGGGCTCTCAGTGCCTCCACCCCAAGCGCCGAGAGAGATGTAC 1099
OY 418 GTCCAACTTTTGGGCAAAATCGACAAAAGCTACATATCAAGAGCAATAGATGGGATC 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1100 GTCCCACTTTCTGGGCTCTCCAGACCAAGATTAATCTACCGTACCCAGAGATGGGACG 1159
OY 478 AATACCAATGCGGCTTGGGGGTGATTCCTGTATCATCATATGATGTGATTCGATT 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1160 AGTTCCCTTAAGCGCGCTGGGCAATTCCTCTTCCCTGGCTTTGGGAGCTGAAGACT 1219
OY 538 ATCAGTTTCAT 547
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1220 ACTACCTCTT 1229

```

RESULT 5

```

US-08-738-000-1
; Sequence 1, Application US/08738000
; Patent No. 6074821
; GENERAL INFORMATION:
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: REDUCTASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: Continental Plaza - 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,000
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NO PCT/CA95/00314
; FILING DATE: 25-MAY-1995
; APPLICATION DATA:
; APPLICATION NUMBER: GB 9410620.0
; FILING DATE: 26-MAY-1994
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1980
US-08-738-000-1

```

Query Match 20.9%; Score 116.4; DB 3; Length 2220;  
 Best Local Similarity 53.3%; Pred. No. 2.1e-27;  
 Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

```
QY 1 TGATGTGACCCCAATTTATATGATGAGATATATCTCAATTTGTGAGCACTGTC 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 TCATCATCAGCAGCGCTTTCTTTGAGGCTGACACATCTTCCGCTTTGTGAGCACTGCA 742
QY 61 GCCAATTTGGAAATAGCTGTCTTGTACCTGGAATTTAGCGCATTAATTAATTCAGG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 CCGACATGGGCATCATCTTGCCCATGCTCCCGGAGATCTTCCCATCCAGGGCTACACT 802
QY 121 GCTTATCCGACATGCTGGTTTGTGCAAAACAATGACGCTGACATATAGGCTGCTT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 803 CCTTCGGACCTTGTGAAGCTGTCCAGCTGAGAGTCCACAGAGATATAGAGACTGCA 862
QY 181 TAGAGCTTATCAGACGACATAGAGCTGTCAAGCTTATGCAATTCACCTGGAGACTG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 863 TTGAGCCATCAAGACAGACGATGTGCTCCGACACTATGAGCTGAGAGCTGTGA 922
QY 241 AATGTGCAAAAAGATTTTACCTCATGGAATTAAGACAT--TGCATCTTTATACACTAA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 923 GCTGTGCGAGAGCTTCTGGCCAGTGTGCTGAGGCTCCACTTCTACACCTCA 982
QY 298 ATATGAGAAATCTGCATTGGCAATTAATGAACTTGGCTTAATTAAGAGTCCAAAG 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 983 ACCCGAGATGGCTTACACAGAGAGTGTGAAGCGCTGGGAGTGTGAGACTGAGACCCA 1042
QY 358 TTTTACGCTCTTACCTTGTGAGAGCGCCCTGCAAAATGTTTCCGTTAATTAAGAGATGTCC 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1043 GCGCTC--CCCTACCTGGGCTCTCAGTCCACCCCAAGCGCCGAGAGAAAGATGTAC 1099
QY 418 GTCCAACTTTTGGGCAATGCAACCAAAAGCTAATTAAGAGACCATAGATGAGATGTC 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 GTCCATCTTCTGGGCTCCAGACCAAAAGTTACATCTACCTGACCCAGAGAGTGGAGC 1159
QY 478 AATACCCACATGGGCTTGGGCTGATCCGTAAATCATCATATGAGCTTATGCTGAT 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1160 AGTTCCTCAAGCGCCGCTGGGCAATTCCTCTTCCCTTGGGAGCTGAGAGACT 1219
QY 538 ATCAGTTCAT 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1220 ACTACCTCTT 1229
```

```
RESULT 6
US-09-258-928-1
; Sequence 1, Application US/09258928
; Patent No. 6218120
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/005002
; CURRENT FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US/09/258,928
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1980)
US-09-258-928-1

Query Match 20.98; Score 116.4; DB 4; Length 2220;
Best Local Similarity 53.38; Pred. No. 2,1e-27;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;
QY 1 TGATGTGACCCCAATTTATATGATGAGATATATCTCAATTTGTGAGCACTGTC 60
```

```
Db 683 TCATCATCAGCAGCGCTTTCTTTGAGGCTGACACATCTTCCGCTTTGTGAGCACTGCA 742
QY 61 GCCAATTTGGAAATAGCTGTCTTGTACCTGGAATTTAGCGCATTAATTAATTCAGG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 CCGACATGGGCATCATCTTGCCCATGCTCCCGGAGATCTTCCCATCCAGGGCTACACT 802
QY 121 GCTTATCCGACATGCTGGTTTGTGCAAAACAATGACGCTGACATATAGGCTGCTT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 803 CCTTCGGACCTTGTGAAGCTGTCCAGCTGAGAGTCCACAGAGATATAGAGACTGCA 862
QY 181 TAGAGCTTATCAGACGACATAGAGCTGTCAAGCTTATGCAATTCACCTGGAGACTG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 863 TTGAGCCATCAAGACAGACGATGTGCTCCGACACTATGAGCTGAGAGCTGTGA 922
QY 241 AATGTGCAAAAAGATTTTACCTCATGGAATTAAGACAT--TGCATCTTTATACACTAA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 923 GCTGTGCGAGAGCTTCTGGCCAGTGTGCTGAGGCTCCACTTCTACACCTCA 982
QY 298 ATATGAGAAATCTGCATTGGCAATTAATGAACTTGGCTTAATTAAGAGTCCAAAG 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 983 ACCCGAGATGGCTTACACAGAGAGTGTGAAGCGCTGGGAGTGTGAGACTGAGACCCA 1042
QY 358 TTTTACGCTCTTACCTTGTGAGAGCGCCCTGCAAAATGTTTCCGTTAATTAAGAGATGTCC 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1043 GCGCTC--CCCTACCTGGGCTCTCAGTCCACCCCAAGCGCCGAGAGAAAGATGTAC 1099
QY 418 GTCCAACTTTTGGGCAATGCAACCAAAAGCTAATTAAGAGACCATAGATGAGATGTC 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 GTCCATCTTCTGGGCTCCAGACCAAAAGTTACATCTACCTGACCCAGAGAGTGGAGC 1159
QY 478 AATACCCACATGGGCTTGGGCTGATCCGTAAATCATCATATGAGCTTATGCTGAT 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1160 AGTTCCTCAAGCGCCGCTGGGCAATTCCTCTTCCCTTGGGAGCTGAGAGACT 1219
QY 538 ATCAGTTCAT 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1220 ACTACCTCTT 1229
```

```
RESULT 7
US-08-961-527-140/c
; Sequence 140, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
```



Mon Feb 24 06:47:48 2003

us-09-720-451-5.rn

Page 6

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9006017.9
FILING DATE: 16-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 215964/SEE35669USCIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4203 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-288-630-3

Query Match
Best Local Similarity 5.3%; Score 35.4; DB 2; Length 4203;
Matches 87; Conservative 0; Mismatches 86; Indels 0; Gaps 0

OY 156 ATACAGCCTACATTTATGCGCTGTAGAGCCTATCAAGACAAATGAAGAGTGTCAAG 215
Db 630 ATGTCAGGTATTAGATTTTGTTCATTAATTAAGGCAACATTTAGACAGGTAAAGA 571
OY 216 GGTATGAGTAATTCACCTCGGACGAAATGTCGCAAAAGATTAGCTCATGAAATAG 275
Db 570 ACTTATGCTCTACTGCTAGACGACGACGACGACGACGACGACGACGACGACGACGAC 511
OY 276 ACATTCGACCTTTATACACTAATATATGAGGAATTCGATTCGCAATCAAT 328
Db 510 TCATGTAATTTCTATGAGAGTGAATGAGTATATCTTGTAAAGACTAT 458

RESULT 10
US-09-221-017B-931
Sequence 931, Application US/09221017B
Patent No 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORBSTER
STREET: 755 PACE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/01023
FILING DATE: 10-DEC-1998

```

```

ATTORNEY/AGENT INFORMATION:
1 NAME: Monroy, Gladys H.
2 REGISTRATION NUMBER: 33,430
3 REFERENCE/DOCKET NUMBER: 27340-20021.00
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: 650-813-5600
6 TELEFAX: 650-494-0792
7 TELEX: 706141
8 INFORMATION FOR SEQ ID NO: 931:
9 SEQUENCE CHARACTERISTICS:
10 LENGTH: 3152 base pairs
11 TYPE: nucleic acid
12 STRANDEDNESS: double
13 MOLECULE TYPE: DNA (genomic)
14 HYPOTHEetical: NO
15 ANTI-SENSE: UNKNOWN
16 ORIGINAL SOURCE:
17 ORGANISM: PORPHYROMONAS GINGIVALIS
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: 1...3152
21 US-09-221-0178-931
22
23 Query Match 5.78; Score 31.8; DB 4; Length 3152;
24 Best Local Similarity 50.3%; Pzed. No. 2.3;
25 Matches 78; Conservativ 0; Mismatches 77; Indels 0; Gaps 0;
26
27 QY 105 ATTAATATTATCAAGGGCTTTATCCGCACTGACTGGTGTTCGAAACAAAGATACCACT 164
28 + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
29 Db 2914 AATGTATACCATATTATTGTTTGTGCGGAATCTGTTGTTTTCGCCGCAAAAGATCCCA 2973
30 + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
31 QY 165 GACATTTATGGCTGCTCTTTAGAGCCATACGACGACATGACGAAAGCTGTGCAAGGCTTATGCA 224
32 + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
33 Db 2974 GAGATTTATCCGATACGACGACGCTTCATGATGCTGTGCGCAAAAGTATAGGAAGATATATC 3033
34 + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
35 QY 225 ATTCACCTGGGAGACTGAAATGTGCACAAAAGATTTT 259
36 + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
37 Db 3034 CACAAAGCGCTGATTTCCTGCACAAAACATTTT 3068
38 + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
39
40 RESULT 11
41 US-09-446-504-41
42 Sequence 41, Application US/09446504
43 Patent No. 6218150
44 GENERAL INFORMATION:
45 APPLICANT: UEMORI, Takashi
46 APPLICANT: SATO, Yoshimi
47 APPLICANT: FUJITA, Tomoko
48 APPLICANT: MIYAKE, Kazuo
49 APPLICANT: MUKAI, Hiroyuki
50 APPLICANT: ASADA, Kiyozo
51 APPLICANT: KATO, Ikunoshi
52 TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
53 FILE REFERENCE: 1432-40BPC
54 CURRENT APPLICATION NUMBER: US/09/446, 504
55 PRIOR FILING DATE: 1999-12-23
56 PRIOR APPLICATION NUMBER: PCT/JP98/02845
57 PRIOR FILING DATE: 1998-06-24
58 PRIOR APPLICATION NUMBER: JP 9-187496
59 PRIOR FILING DATE: 1997-06-26
60 PRIOR APPLICATION NUMBER: JP 9-320692
61 PRIOR FILING DATE: 1997-11-27
62 NUMBER OF SEQ ID NOS: 92
63 SOFTWARE: Patentin Ver. 2.1
64 SEQ ID NO 41
65 LENGTH: 989
66 TYPE: DNA
67 ORGANISM: Artificial Sequence
68 FEATURE:
69 OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
70 US-09-446-504-41

```



## CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourile and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/238,163  
FILING DATE: 03-MAY-1994

## CLASSIFICATION: 800

## ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 2307E-540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2075 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

## FEATURE:

NAME/KEY: CDS  
LOCATION: 421..1401  
US-08-238-163-3

Query Match 5.6%; Score 31.2; DB 1; Length 2075;

Best Local Similarity 45.5%; Pred. No. 2.9; Mismatches 133; Indels 0; Gaps 0;

Matches 111; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 123 TTTATCCGATGACTGGTGGTTTTCACAAACAAAGATACACGCTGACATTATGCGCTTAA 182  
DB 340 TATTTAGCAATAGCGATTGTTGGTCAGAAAAGAACTGCCACAAAGTACATGACATG 281  
QY 183 GAGCCTATCAAGACATGATAGAGAGCTGCAAGCTGTATGAAATTCACCTGGAAGTGA 242  
DB 280 GTGTGATCAGATCTAGATATTAATGCGTGGCTTTGAAATTCACCTGCAAAAGAG 221  
QY 243 ATGTCAAAAAGATTTAGCTCATGCAATTAAGACATTGATCTTTATACATAAATATG 302  
DB 220 ATGTCAATGATTTTACAAATTTATATCATGATGTTTATATCATAGATATTTAATTA 161  
QY 303 GAGAAATCTGCATTTGCATTAATGACCTTGCGCTAATTTGAAGAGTCCAAAGTTCT 362  
DB 160 AATGATATATAAATAAAGATACATAAATAAATGTTTAAATGTTAAGTCTTT 101  
QY 363 AGGT 366  
DB 100 AAGT 97

## RESULT 15

US-09-592-891A-13/C

Sequence 13. Application US/09592891A

Patent No. 6329174

GENERAL INFORMATION:

APPLICANT: Xiao-Zhou Michelle Wang

APPLICANT: Xavier Georges Sarda

APPLICANT: Michael David Tomalski

APPLICANT: Vincent Paul Mary Wingate

TITLE OF INVENTION: Helicobis Glutamate Receptor

FILE REFERENCE: A32815 07267 0118

CURRENT APPLICATION NUMBER: US/09/592,891A

CURRENT FILING DATE: 2001-01-22

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO: 13

LENGTH: 4621

TYPE: DNA

ORGANISM: Helicobis virescens

FEATURE:

NAME/KEY: unsure

LOCATION: (764)...(764)

OTHER INFORMATION: n is a, t, g, or c

US-09-592-891A-13

Query Match 5.6%; Score 31; DB 4; Length 4621;

Best Local Similarity 51.9%; Pred. No. 4.9; Mismatches 65; Indels 0; Gaps 0;

Matches 70; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 241 AAATGTCAAAAAGATTTTACCTCATGCAATTAAGACATTGATCTTTATACATAATA 300  
DB 2041 AAATATATCAAAAATTTTAAATCTCGAAAAGTAAATATTTTGTACCAAAATTA 1982  
QY 301 TGAGAAATCTGCATTTGCATTAATGAAACCTTGCGCTAATTTGAAGAGTCCAAAGTTT 360  
DB 1981 TTGAGACCTTTATAGTGCAGATTAAGAGTTCTGCTTTAGATGACCAATCAAAATTTGA 1922  
QY 361 CTAGTCTTACCTT 375  
DB 1921 GTCTATTTTACCCT 1907

Search completed: February 12, 2003, 15:12:59

Job time : 132 secs



QY 181 TAGAGCTATCAAGCAATGAGAGCTGTCAAGCTTATGGAATTCACCTGGGAACGTG 240  
DB 808 TAGAGCTATCAAGCAATGAGAGCTGTCAAGCTTATGGAATTCACCTGGGAACGTG 867  
QY 241 AAATGTGCAAAAAGATTTTACCTATGGAATTTAAGACATTCATTTTATACATAAATA 300  
DB 866 AAATGTGCAAAAAGATTTTACCTATGGAATTTAAGACATTCATTTTATACATAAATA 927  
QY 301 TGGAGAAATCGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 360  
DB 928 TGGAGAAATCGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 987  
QY 361 CTAGTCCTTACCTGAG 420  
DB 988 CTAGTCCTTACCTGAG 1047  
QY 421 CAATCTTTGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 480  
DB 1048 CAATCTTTGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1107  
QY 481 ACCCAATGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGG 540  
DB 1108 TCCCAATGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGG 1167  
QY 541 AGTTCATGCGGCA 554  
DB 1168 AGTTCATGCGGCA 1181

## RESULT 2

US-09-734-017A-61  
; Sequence 61, Application US/09734017A  
; Patent No. US20020142422A1  
; GENERAL INFORMATION:  
; APPLICANT: Lerchl, Jens  
; APPLICANT: Renz, Andreas  
; APPLICANT: Ehardt, Thomas  
; APPLICANT: Reinol, Andreas  
; APPLICANT: Clippus, Petra  
; APPLICANT: Bischoff, Friedrich  
; APPLICANT: Frank, Markus  
; APPLICANT: Freund, Annette  
; APPLICANT: Duwendt, Elke  
; APPLICANT: Schmidt, Ralf-Michael  
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved  
; TITLE OF INVENTION: the  
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides and  
; FILE REFERENCE: BASF-NAE-1331-99-US  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR FILING DATE: 1998-12-16  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: Patentin Ver. 2.1/WordPerfect  
; SEQ ID NO 61  
; LENGTH: 574  
; TYPE: DNA  
; ORGANISM: Physcomitrella patens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(573)  
; OTHER INFORMATION: 86\_pprot1\_094\_g10  
US-09-734-017A-61

Query Match 49.0%; Score 273.4; DB 10; Length 574;  
Best Local Similarity 69.7%; Pred. No. 1.2e-69;  
Matches 384; Conservative 0; Mismatches 166; Indels 1; Gaps 1;  
QY 1 TGATGTGACCCCAATATTTATGATAGGATATATCTTCAATATTTGAGAGAGAGAGAG 60  
DB 8 TCATATACACCAGCTGTTATGATACGATATCTTTTGAATTTGATGATGATGATGATG 67

QY 61 GCCAAATGGAATTAACGATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 120  
DB 68 GTCAAAATGGAATTAACGATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 127  
QY 121 GCTTATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
DB 128 GCTTATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 187  
QY 181 TAGAGCTATCAAGCAATGAGAGCTGTCAAGCTTATGGAATTCACCTGGGAACGTG 240  
DB 188 TAGAGCTATCAAGCAATGAGAGCTGTCAAGCTTATGGAATTCACCTGGGAACGTG 247  
QY 241 AAATGTGCAAAAAGATTTTACCTATGGAATTTAAGACATTCATTTTATACATAAATA 300  
DB 248 AAATGTGCAAAAAGATTTTACCTATGGAATTTAAGACATTCATTTTATACATAAATA 307  
QY 301 TGGAGAAATCGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 360  
DB 308 TGGAGAAATCGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 367  
QY 361 CTAGTCCTTACCTGAG 420  
DB 368 CTAGTCCTTACCTGAG 427  
QY 421 CAATCTTTGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 480  
DB 428 CAATCTTTGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 487  
QY 481 ACCCAATGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGG 540  
DB 488 TCCCAATGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGG 546  
QY 541 AGTTCATGCGG 551  
DB 547 AGTTCATGCGG 557

## RESULT 3

US-09-923-876-5596  
; Sequence 5596, Application US/09923876  
; Patent No. US2002013958A1  
; GENERAL INFORMATION:  
; APPLICANT: Laigudi, Raghunath V.  
; APPLICANT: Kamigaki, Laura Y. (Ito)  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING  
; FILE REFERENCE: PL-0012-1.COM  
; CURRENT FILING DATE: 2001-08-06  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR FILING DATE: 1998-05-05  
; NUMBER OF SEQ ID NOS: 6332  
; SOFTWARE: PERL Program  
; SEQ ID NO 5596  
; LENGTH: 264  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457310H1  
; LOCATION: 81..183  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-923-876-5596

Query Match 22.3%; Score 124.4; DB 10; Length 264;  
Best Local Similarity 75.5%; Pred. No. 1.2e-26;  
Matches 166; Conservative 0; Mismatches 53; Indels 1; Gaps 1;  
QY 340 TAATGAGAGTCCAAAGTTCTAGTCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 399

```

Db      4  TAATTGAGGATGCGCAAGGTTTCAAGGCCATTACCTTGGAGGCGCAGCAGCACTAATGTTTCC 63
Qy      400 GTGTTAAAGAGATG-TCCGTCCATCTTTGGGCAATTCACCAAAAACCTACATATCA 458
Db      64  GTGTAAAGAGATGTTMCAGACCTATATTTCTGGGCCAACAACCAAGACCTATCTTAA 123
Qy      459 AGAGCAATGATGGATGATATCCAGATGGGCGCTGGGGGATTCCTGATATCATCA 518
Db      124 AGNCATTTAGTTGGGATCAGTATCCCATGACCGTGGGTGATTTCTGGAAACCATCN 183
Qy      519 TATGTGCAATTAATCTGATTAATCAGTTGATGCGGCCACGTC 558
Db      184 CATGAGACACTTACTGACCAACCACTTCACAAAGACCAAG 223

RESULT 4
US-09-931-795-3
; Sequence 3, Application US/09931795
; Publication No. US20020198211A1
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/005003
; CURRENT APPLICATION NUMBER: US/09/931,795
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 09/592,595
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/738,000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: PCT/CA95/00314
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)...(1983)
US-09-931-795-3

Query Match      20.9%; Score 116.4; DB 9; Length 2219;
Best Local Similarity 53.3%; Pred. No. 5.9e-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

Qy      1  TGATTTGACCCCAATTTATATATGATACGATATATCTCTCAAAATTTGTGAAGACTGTC 60
Db      683 TCATCATACGACGCTTCTTGTGAGCTGCACACATTTCTCCGCTTGTGAAGAGCA 742
Qy      61  GCCAAATTTGGAATACGTGCTCTATTGTACCTGGAATTAATGCGCAATTAATTTACAG 120
Db      743 CCGACATGGGCACTGCTGCCCCATGCTCCCGGGATTTCTTCCATCCAGGGCTACACT 802
Qy      121 GCTTATCCGATGACTGGGTTTGGCAAAACAAGTATACCTGACCTGAGCTGCTT 180
Db      803 CCGTTCGGGAGCTTGTGAAGCTGTCCAAAGCTGGAGGCTGCACAGAGATCAAGAGCTGA 862
Qy      181 TAGAGCCTATCAAGCAATGAAGAAGCTGTCAAGGCTTATGAAATTCACCTGGAGAGT 240
Db      863 TTGAGCCCAATCAAGACAAGATGCTGCATCCGCAATGATGAGCTGAGCTGAGTGA 922
Qy      241 AAATGTGCAAAAAGATTTTACCTATGATTAAGCAT--TGCACTTTTATACCTAA 297
Db      923 GCGTGTCCAGAGGCTTGTGGCAGAGTGGCTTGGTGCAGAGGCTCCACACTTCAACCTCA 982
Qy      298 ATATGGAATATCTGATTTGGCAATTAATGAACTTGGCTTAATTAAGAGTCCAAAG 357

```

```

Db      983 ACCCGAGATGCTTACACACAGAGGTGCTGAAGCCGCTGGGGAATGTGACTGAGACCCA 1042
Qy      358 TTTCATAGTCTCTTACCTTTGGAGAGAGCCCTGCAAAATGTTTCCGTGTTAAAGATGTC 417
Db      1043 GGGGTC--CCCTACCTGGGCTGTGAGGCCACCCCAAGCCGAGAGGAAATGATGAC 1099
Qy      418 GTCCATCTTTTGGGCAAAATGCAACCAAAAGCTACATATCAAGAGACCATGAGAGGATC 477
Db      1100 GTCCATCTTTGCGGCGCTCCAGACCAAAAGATTAATCACTACCATCCAGAGTGGAGC 1159
Qy      478 AATACCAATGAGCGCTTGGGATGATCTCTGAATTCATATATGTCATATCTGATT 537
Db      1160 AGTCCCTTACAGGCGCTGGGCAATTCCTCTCCCTTGGGAGCTGAAGACT 1219
Qy      538 ATCACTTCAT 547
Db      1220 ACTACCTCTT 1229

RESULT 5
US-09-728-910-3
; Sequence 3, Application US/09728910
; Patent No. US20010025030A1
; GENERAL INFORMATION:
; APPLICANT: Rozen, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/006001
; CURRENT APPLICATION NUMBER: US/09/728,910
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)...(1983)
US-09-728-910-3

Query Match      20.9%; Score 116.4; DB 10; Length 2219;
Best Local Similarity 53.3%; Pred. No. 5.9e-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

Qy      1  TGATTTGACCCCAATTTATATATGATACGATATATCTCTCAAAATTTGTGAAGACTGTC 60
Db      683 TCATCATACGACGCTTCTTGTGAGCTGCACACATTTCTCCGCTTGTGAAGAGCTCA 742
Qy      61  GCCAAATTTGGAATACGTGCTCTATTGTACCTGGAATTAATGCGCAATTAATTTACAG 120
Db      743 CCGACATGGGCACTGCTGCCCCATGCTCCCGGGATTTCTTCCATCCAGGGCTACACT 802
Qy      121 GCTTATCCGATGACTGGGTTTGGCAAAACAAGATACAGCTGACCTGATATGCTGCTT 180
Db      803 CCGTTCGGGAGCTTGTGAAGCTGTCCAAAGCTGGAGGCTGCACAGAGATCAAGAGCTGA 862
Qy      181 TAGAGCCTATCAAGCAATGAAGAAGCTGTCAAGGCTTATGAAATTCACCTGGAGAGT 240
Db      863 TTGAGCCCAATCAAGACAAGATGCTGCATCCGCAATGATGAGCTGAGCTGAGTGA 922
Qy      241 AAATGTGCAAAAAGATTTTACCTATGATTAAGCAT--TGCACTTTTATACCTAA 297
Db      923 GCGTGTCCAGAGGCTTGTGGCAGAGTGGCTTGGTGCAGAGGCTCCACACTTCAACCTCA 982
Qy      298 ATATGGAATATCTGATTTGGCAATTAATGAACTTGGCTTAATTAAGAGTCCAAAG 357
Db      983 ACCCGAGATGCTTACACAGAGGTGCTGAAGCCGCTGGGGAATGTGACTGAGACCCA 1042

```

```
OY 358 TTCTAGGTCCTTACCTTGAGAGACCGCTCGCAATGTTTTCGCTGTTAAAGATGTCC 417
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1043 GCGGTC---CCCTACCTGGGCGCTTCAGTGGCCACCCCAAGCGCGGAGAGAAATGTCC 1099
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 418 GTCCAACTCTTTGGGCAAAATCGCAAAAAGCTACATATCAAGCAATAGATGGATC 477
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1100 GTCCAACTCTTTGGGCGCTTCAGAGCAAAAGTTCATCTACCGTACCGAGAGTGGAGC 1159
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 478 AATACCCATGCGCGCTGGGGGTATCTCTATTCATCATATATGTCATATATCTATT 537
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1160 AATTCCTTACGCGCGCTGGGGCAATTCCTCTTCCCTGCTTTGGAGAGCTGAAGACT 1219
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 538 ATCAGTTTCAT 547
      | | | | |
Db 1220 ACTACCTCTT 1229

RESULT 6
US-09-931-795-1
; Sequence 1, Application US/09931795
; Publication No. US20020198211A1
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETERAHYDROFOLATE
; FILE REFERENCE: 04844/005003
; CURRENT FILING DATE: US/09/931,795
; PRIOR APPLICATION NUMBER: US 09/592,595
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/738,000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: PCT/CA95/00314
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1980)
US-09-931-795-1
```

Query Match 20.9%; Score 116.4; DB 9; Length 2220;  
Best Local Similarity 53.3%; Pred. No. 5.9e-24;  
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

```
OY 1 TGATGTGACCCCAATTTATTTATGATGAGATATATCTCAATTTGTGAAGACGTGC 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 TCATCATACGAGAGCTTTCTTTGAGGCTGACACATCTTCGCTTGTGAAGCATGCA 742
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 61 GCCAAATGGAATAGGTGCTTATTTACTGGAATTTAGCCCATTAATTAATTAAGG 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 743 CCGAATGAGCATCACTTCCCATCGTCCCGGAGTTTCCATCCAGGCTACCACT 802
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 121 GCTTATCCGATGACTGGTTTTCGAAAACAAGATACCAGCTACATTATGCGTCTT 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 803 CCTTTGGAGCTTTGGAAGCTGTCCAGCTGAGGAGTGCCACAGAGATTCAGAGAGTGA 862
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 181 TAGAGCTTACAGACAAATGAAGAGCTGTCAAGCTTATGGAATTCACCTGGAGATG 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 863 TTGAGCCCAATCAAGACAAAGATGCTGCATCCGCAACTATGGCATTCAGTGGCGGTGA 922
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 241 AATGTGCAAAAAGATTTTATGCTCATGGAATTAAGACAT--TGCATCTTTATACATTA 297
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 923 GCGTGTGCGAGAGCTTTGCGCAAGTGGCTTGGTCCAGGCTCCACTCTTACACCTCA 982
```

```
OY 298 ATATGAGAAATCTGATATGAGCAATACATATGACCTTGGCTAATTTGAAGATCCAAAG 357
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 983 ACGGCAATGAGCTTACCAAGAGAGTGTGATGAGGATGTGACTGAGAGACCCCA 1042
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 358 TTCTAGGTCCTTACCTTGAGAGACCGCTCGCAATGTTTTCGCTGTTAAAGATGTCC 417
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1043 GCGGTC---CCCTACCTGGGCGCTTCAGTGGCCACCCCAAGCGCGGAGAGAAATGTCC 1099
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 418 GTCCAACTCTTTGGGCAAAATCGCAAAAAGCTACATATCAAGCAATAGATGGATC 477
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1100 GTCCAACTCTTTGGGCGCTTCAGAGCAAAAGTTCATCTACCGTACCGAGAGTGGAGC 1159
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 478 AATACCCATGCGCGCTGGGGGTATCTCTGTAATTCATCATATATGTCATATATCTATT 537
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1160 AATTCCTTACGCGCGCTGGGGCAATTCCTCTTCCCTGCTTTGGAGAGCTGAAGACT 1219
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 538 ATCAGTTTCAT 547
      | | | | |
Db 1220 ACTACCTCTT 1229

RESULT 7
US-09-728-910-1
; Sequence 1, Application US/09728910
; Patent No. US20010025030A1
; GENERAL INFORMATION:
; APPLICANT: Rozen, Jaspereet
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETERAHYDROFOLATE
; FILE REFERENCE: 04844/006001
; CURRENT FILING DATE: US/09/728,910
; PRIOR APPLICATION NUMBER: US 09/12-01
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1980)
US-09-728-910-1
```

Query Match 20.9%; Score 116.4; DB 10; Length 2220;  
Best Local Similarity 53.3%; Pred. No. 5.9e-24;  
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

```
OY 1 TGATGTGACCCCAATTTATTTATGATGAGATATATCTCAATTTGTGAAGACGTGC 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 TCATCATACGAGAGCTTTCTTTGAGGCTGACACATCTTCGCTTGTGAAGCATGCA 742
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 61 GCCAAATGGAATAGGTGCTTATTTACTGGAATTTAGCCCATTAATTAATTAAGG 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 743 CCGAATGAGCATCACTTCCCATCGTCCCGGAGTTTCCATCCAGGCTACCACT 802
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 121 GCTTATCCGATGACTGGTTTTCGAAAACAAGATACCAGCTACATTATGCGTCTT 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 803 CCTTTGGAGCTTTGGAAGCTGTCCAGCTGAGGAGTGCCACAGAGATTCAGAGAGTGA 862
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 181 TAGAGCTTACAGACAAATGAAGAGCTGTCAAGCTTATGGAATTCACCTGGAGATG 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 863 TTGAGCCCAATCAAGACAAAGATGCTGCATCCGCAACTATGGCATTCAGTGGCGGTGA 922
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 241 AATGTGCAAAAAGATTTTATGCTCATGGAATTAAGACAT--TGCATCTTTATACATTA 297
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 923 GCGTGTGCGAGAGCTTTGCGCAAGTGGCTTGGTCCAGGCTCCACTCTTACACCTCA 982
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 298 ATATGAGAAATCTGATATGAGCAATACATATGACCTTGGCTAATTTGAAGATCCAAAG 357
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 983 ACGGCAATGAGCTTACCAAGAGAGTGTGATGAGGATGTGACTGAGAGACCCCA 1042
```

```
OY 358 TTCTAGTCTTACCTTGAGAGCCCTGCAATGTTTCCGTGTAAGAAGATGTC 417
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1043 GGGCTC---CCCTACCTGGGCTCTCAAGGCCCAAGCCGAGAGAAATGTTAC 1099
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 418 GTCCATCTTTTGGGCAATGACCAAAAAGCTACATATCAAGACCATAGATGGATC 477
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1100 GTCCATCTTTGCGGCTCCAGACCAAAAGTTATCTACCTACCCAGAGTGGGACG 1159
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 478 AATACCAATGAGGCGTTGGGCTGATTCCTGTAATCATATATAGTGCTATTCGATT 537
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1160 AGTTCCCTTAACGCGCGCTGGGCAATTCCTCCCTGCTTGGGAGCTGAAGAGCT 1219
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 538 ATCAGTTGAT 547
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1220 ACTACTCTT 1229
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 8

```
US-09-923-876-5753
; Sequence 5753, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Itco)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SPEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/2298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/0085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 8332
; SOFTWARE: PERL Program
; SEQ ID NO 5753
; LENGTH: 181
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457642H1
; NAME/KEY: unsure
; LOCATION: 2, 165, 173, 180
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5753
```

## Query Match

```
Best Local Similarity 20.2%; Score 112.6; DB 10; Length 181;
Best Local Similarity 76.4%; Pred. No. 2.5e-23;
Matches 136; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```
OY 337 GCGTAATTGAGAGTCCAAAGTTTCTAGTCTTACCTTGAGAGCGCCCTGCAATGTTT 396
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 GATTAATTGAGAGTCCAAAGTTTCAAGGCTTACCTTGAGAGCGCCGCAATGTTT 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 397 TCCGTTTAAAGAGATGTCGCTCATCTTTGGGCAATCGCAAAAAGCTCATAT 456
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 TCCGTTTAAAGAGATGTTGACCTATATTTCGAGCCCAAGACCAAGAGCTATCTTA 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 457 CAAGGACCATGAGTGGATCATATACCATAGGAGTGGGCTGCTGTAATTC 514
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 AAGGACATTAGTTGGATCAGTATCCCATGAGCGGTGGGATTCGMAACC 178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 9

```
US-09-294-093B-4653
; Sequence 4653, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
```

```
;; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
;; FILE REFERENCE: PL-0009 US
;; CURRENT APPLICATION NUMBER: US/09/294,093B
;; CURRENT FILING DATE: 1999-04-16
;; PRIOR APPLICATION NUMBER: 60/082,567
;; PRIOR FILING DATE: April 21, 1998
;; NUMBER OF SEQ ID NOS: 6207
;; SOFTWARE: PERL Program
;; SEQ ID NO 4653
;; LENGTH: 159
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: misc_feature
;; OTHER INFORMATION: Incyte ID No. US20010051335A1 7003547A1H1
;; NAME/KEY: unsure
;; LOCATION: 14
;; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4653
```

## Query Match

```
Best Local Similarity 17.4%; Score 97; DB 10; Length 159;
Best Local Similarity 76.6%; Pred. No. 7.9e-19;
Matches 118; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
```

```
OY 388 CAATGTTTCCGTGTPAAGAGATGTCCTCCATCTTTTGGGCAATGACCAAAA 447
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6 CTATGTTTCCGTGTPAAGAGATGTTGACCTATATTCTGGGCCAAGACCAAGA 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 448 GCTACATATCAAGSCATAGATGGATCATATCCCATGAGCGCTGGGATTCCT 507
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 66 GCTATCTTAAAGACATATAGGTGGATCATATCCCATGAGCGGTGGATTCCT 125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 508 GCAATCATATAGTGGATCATATCTGATTATCA 541
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 126 GGAACCATCATATGAGACACTTACTGACCAACA 159
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 10

```
US-09-734-017A-63
; Sequence 63, Application US/09734017A
; Patent No. US2002014242A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reinold, Andreas
; APPLICANT: Clipus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Anneke
; APPLICANT: Duwendig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides an
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/171,100
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/Mordperfect
; SEQ ID NO 63
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Physcomitrella patens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(409)
; OTHER INFORMATION: 62_mm20_c10rev
US-09-734-017A-63
```



```

; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 1
; LENGTH: 127197
; TYPE: DNA
; ORGANISM: glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 515002_region_g2
US-09-754-853A-1

```

```

Query Match          7.2%; Score 40; DB 9; Length 127197;
Best Local Similarity 55.9%; Pred. No. 0.43;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

```

```

QY 193 AGACACATGAGAGAGCTGTCAAGCTTATGGAATTCACCTGGAGCAATGTCAGAAA 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58616 AGGGGAATGATGAGAGCTGCAGCCCTAAGAAAAGATTTCACCTTGTCAAAAAGGGGAGCA 58557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 253 AGATTTTTCATGAGATTAAGACATTCATCTTATACCTTAATATGAGAAATCTG 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58556 AGAAGGCTCAAAATGATGATGAGATCTCATCATGATACCTTATCAAGAGGTACAG 58497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 313 CATGCGCAATCTAAT 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58496 CATCTGGAACAGAAAT 58481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
US-09-815-242-7047
; Sequence 7047, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7047
; LENGTH: 1266

```

```

; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1266)
US-09-815-242-7047

```

```

Query Match          6.2%; Score 34.8; DB 10; Length 1266;
Best Local Similarity 60.6%; Pred. No. 1.9;
Matches 57; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

QY 353 CAAAGTTCTAGGTCCTTACTTGGAGAGCCCTGCAATGTTTCCGTGAAGAAGA 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 411 CAAATTTATTAACCTCGTACTTTATGGAATTAAGTCTGCGATGCTTAATGATGAGA 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 TGTGCTCAATCTTTGGCAATGACCAANA 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 471 TGTGCTCAAAAGCATGAGATGTAACCAANA 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-09-938-842A-4584/C
; Sequence 4584, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING.
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4584
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4584

```

```

Query Match          6.0%; Score 33.4; DB 9; Length 2000;
Best Local Similarity 49.2%; Pred. No. 5.9;
Matches 88; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```

```

QY 273 AAGACATTCATCTTATACCTAATATGAGAAATCGCATTTGGCAATCTAATGAC 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 755 AAGACATTAACATTTTAAAGTTAGATGAGTTTTCGAAAGTGAAGAAAAAATCAAA 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 333 CTGGCGTAATGAGAGTCCAAAGTTCTAGTCTTACCTTGGAGAGCCCTGCAAT 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 695 ATTATTTTAAATCAATTAATCCAAATCATTAATCCAAACGATGCCACCTAAT 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 393 GTTTCGGTGAAGAGATGTCGCAATCTTTGGCAATGACCAAAAAAGCTA 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 635 ATTAACTATTAATAAGGTTTGAAGCAATGATATATAAAGCTAAAAAATATA 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: February 23, 2003, 19:39:34  
Job time : 240 secs

7  
.  
.

Mon Feb 24 06:47:51 2003

us-09-720-451-5.rst

Page 1

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 13:48:04 : Search time 2194 Seconds  
(without alignments)  
4118.998 Million cell updates/sec

Title: US-09-720-451-5  
Perfect score: 558  
Sequence: 1 tgaattaccacattattt.....tcaagtcacgcgcacagcg 558

Scoring table: IDENTITY-MUC  
Gapex 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estnu:\*  
4: em\_estov:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estro:\*  
9: gp\_est1:\*  
10: gp\_est2:\*  
11: gp\_est3:\*  
12: gp\_est4:\*  
13: gp\_est4:\*  
14: gp\_est5:\*  
15: em\_estfun:\*  
16: em\_estfun:\*  
17: gp\_est1:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556.4	99.7	588	12	BF071278 st42e03.y
2	523.6	93.8	565	9	AT900449 sc07a01.y
3	484.4	86.8	759	13	BT308605 EST50015
4	482.8	86.5	755	13	BT308302 EST529712
5	466.4	83.6	573	10	AM203249 st72711.y
6	466.4	83.6	579	10	AM203207 st72702.y

7	429.4	77.0	468	12	BG405762	BG405762 sac37108.
8	422	75.6	668	14	BQ404086	BQ404086 GA_Ed006
9	400.4	71.8	589	12	BF269445	BF269445 GA_ED000
10	399.2	71.5	593	10	BE204232	BE204232 EST396908
11	378	67.7	558	10	BE433609	BE433609 EST400138
12	376.6	67.5	556	10	BE433601	BE433601 EST400130
13	369.2	66.2	661	14	BQ863196	BQ863196 GGC23015.
14	366.8	65.7	652	12	BC904512	BC904512 Tair1132G
15	364.6	65.3	560	13	BMS26290	BMS26290 sal39404.
16	362.2	64.9	592	10	BE354491	BE354491 EST355834
17	358	64.2	2423	11	AT109096	AT109096 Zea_mays
18	357.6	64.1	569	14	BM885389	BM885389 sal199b11.
19	351	62.9	657	14	BQ471589	BQ471589 HV020031
20	348.8	62.5	419	10	AM234903	AM234903 sf120e05.y
21	344.6	61.8	568	9	AT729106	AT729106 BMLGH1126
22	338.8	60.7	505	10	AM029711	AM029711 EST72966
23	338.6	60.7	505	14	BQ246637	BQ246637 Tair15007F
24	337.8	60.5	516	10	AM649374	AM649374 EST327828
25	335	60.0	582	10	AM931213	AM931213 EST357056
26	333	59.7	397	12	BR070260	BR070260 st14b03.y
27	332.2	59.5	523	10	BE403916	BE403916 WHE0415_G
28	329	59.0	530	13	B1123572	B1123572 I063P12P
29	328.2	58.8	627	10	AV915401	AV915401 AV915401
30	326	58.4	595	14	BQ471487	BQ471487 HV02J07r
31	322.6	57.8	678	9	AT727725	AT727725 BMLGH1883
32	317.4	56.9	685	13	B1921517	B1921517 EST541420
33	317	56.8	568	10	BE346411	BE346411 sp25c05.y
34	317	56.8	634	10	BE346389	BE346389 sp25a05.y
35	313.8	55.2	538	13	B1131003	B1131003 G113P2Y
36	311.8	55.9	642	14	BQ862545	BQ862545 GGC21606
37	311.4	55.8	814	10	BE194793	BE194793 HVCEB000
38	308.2	55.2	438	13	B1129696	B1129696 C097P07Y
39	302.6	54.7	760	14	BQ294876	BQ294876 OGE8G21.y
40	302.6	54.2	548	13	B1921139	B1921139 EST941042
41	289.4	51.9	511	13	B1131121	B1131121 G115P70Y
42	286.6	51.6	613	14	BQ402630	BQ402630 GA_Ed005
43	286.6	51.4	535	10	BE215671	BE215671 HV_CEB000
44	283	50.7	387	13	B1127775	B1127775 G065P81Y
45	281.6	50.5	700	10	BE576981	BE576981 LO-1588T3

## ALIGNMENTS

RESULT 1	BF071278	588 bp	mRNA	linear	EST 06-DEC-2001
LOCUS	st42e03.y1	Gm-c1067	glycine max	cDNA clone	GENOME SYSTEMS CLONE ID:
DEFINITION	Gm-c1067-1734 5' similar to SW:MPHR_ARATH_O80585 PROBABLE				
ACCESSION	BF071278	GI:10846202			
VERSION	BF071278				
KEYWORDS	EST.				
SOURCE	soybean.				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
REFERENCE	1 (bases 1 to 588)				
AUTHORS	Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Corvett,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.				
TITLE	Unpublished (1999)				
JOURNAL	Public Soybean EST Project				
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800				

Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: curesgen.com  
Insert Length: 1779 Std Error: 0.00  
High quality sequence stop: 414.

## FEATURES

source

1. 588  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl0107-1734"  
/clone\_lib="Gm-cl067"  
/tissue\_type="germinating shoot, 3 day old seedling, auxin  
treatment  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from germinating shoots of 3 day old seedling for the  
cultivar Williams 82. The seedlings were germinated in a  
growth chamber using germination paper in a solution  
containing 100ppm auxin. Complementary DNA was synthesized  
from mRNA using a primer consisting of a poly(dT) sequence  
with a XhoI restriction site. EcoRI adapters were ligated  
to the blunt-ended cDNA fragments followed by XhoI  
digestion. The cDNA fragments were directionally cloned  
into the EcoRI-XhoI restriction site of the pBluescript  
vector. The ligated cDNA fragments were transformed into  
DH10B host cells (GibcoBRL). This library was constructed  
in the laboratory of Dr. Randy Shoemaker."

BASE COUNT

182 a 114 c 124 g 168 t

ORIGIN

Query Match 99.7%; Score 556.4; DB 12; Length 588;  
Best Local Similarity 99.8%; Pred. No. 6.6e-153;  
Matches 557; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGATGTGACCCATTTTATGATGAGGATATTCCTCAATTTTGACGACTGTC 60  
DB 5 TGATGTGACCCATTTTATGATGAGGATATTCCTCAATTTTGACGACTGTC 64  
QY 61 GCCAATGTGAATACGTGCTTATGCTGATGATGATGATGATGATGATGATG 120  
DB 65 GCCAATGTGAATACGTGCTTATGCTGATGATGATGATGATGATGATGATG 124  
QY 121 GCTTATCCGATGCTGCTTATGCTGATGATGATGATGATGATGATGATG 180  
DB 125 GCTTATCCGATGCTGCTTATGCTGATGATGATGATGATGATGATGATG 184  
QY 181 TAGAGCCATGAGGATGATGATGATGATGATGATGATGATGATGATG 240  
DB 185 TAGAGCCATGAGGATGATGATGATGATGATGATGATGATGATGATG 244  
QY 241 AATGTGCAAAAGATTTAGCTATGATGATGATGATGATGATGATGATGATG 300  
DB 245 AATGTGCAAAAGATTTAGCTATGATGATGATGATGATGATGATGATGATG 304  
QY 301 TGGAGCAATGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 360  
DB 305 TGGAGCAATGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 364  
QY 361 CTAGTCTTACCTTGGAGAGCCCTGCAATGTTTCCGTTAAAGAGATGTCGTC 420  
DB 365 CTAGTCTTACCTTGGAGAGCCCTGCAATGTTTCCGTTAAAGAGATGTCGTC 424  
QY 421 CAATCTTTTGGGCAAAATGCAACAAAAGCTACATCAAGAGCATTAGATGGATCAAT 480  
DB 425 CAATCTTTTGGGCAAAATGCAACAAAAGCTACATCAAGAGCATTAGATGGATCAAT 484  
QY 481 ACCCAATGGGCGTTGGGGGATGCTGATGATGATGATGATGATGATGATGATG 540  
DB 485 ACCCAATGGGCGTTGGGGGATGCTGATGATGATGATGATGATGATGATGATG 544

QY 541 AGTCATGCGGCACGCTG 558  
DB 545 AGTCATGCGGCACGCTG 562

## RESULT 2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 565)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khama

A., Bolla, B., Marr, M., Hillier, L., Kueba, T., Martin, J., Beck, C.,

Wylie, T., Underwood, K., Steptoe, M., Rheising, B., Allen, M., Bowers

J., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck

R., Ritter, E., Kohn, S., Shin, T., Jackson, T., Cardenas, M., Mccann

R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

Possible reversed clone: similarity on wrong strand This clone is

available through: Resgen, Invitrogen Corp. 2130 South Memorial

Parkway Huntsville, AL 35801 For further information call: (800

)-533-4363 or contact via email: curesgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 401.

Location/Qualifiers

1. 565

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl012-1441"

/clone\_lib="Gm-cl012"

/tissue\_type="Apical shoot tips, 9-10 day old etiolated

seedlings

/lab\_host="XL10-Gold"

/note="Vector: pBluescript II Xr; Site\_1: EcoRI; Site\_2:

XhoI; This cDNA library was constructed from mRNA isolated

from the apical shoots of 9 to 10 day old etiolated

seedlings. The shoot tips including any emerged leaves

were harvested for mRNA isolation. The cDNA library was

prepared using the Stragene pBluescript II Xr cDNA

library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a poly

(dT) sequence with a XhoI restriction site. EcoRI adapters

were ligated to the blunt-ended cDNA fragments followed by

XhoI digestion. The cDNA fragments were directionally

cloned into the EcoRI-XhoI restriction site of the

pBluescript vector. The ligated cDNA fragments were

transformed into XL10-Gold host cells. This library was

constructed by Dr. Randy Shoemaker and Dr. John

Erpelting."

BASE COUNT

171 a 112 c 117 g 163 t 2 others

ORIGIN

Query Match 93.8%; Score 523.6; DB 9; Length 565;  
Best Local Similarity 97.3%; Pred. No. 2.9e-143;  
Matches 532; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

[illegible]

LOCUS	DEFINITION	BI308302	755 bp	MRNA	linear	EST 20-JUL-2001
LOCUS	EST29712 GP0D Medicago truncatula cDNA clone pGP0D-2P15 5' end,					
ORIGIN	1. 759					
	/organism="Medicago truncatula"					
	/cultivar="A17"					
	/db_xref="taxon:3880"					
	/clone="pGP0D-714"					
	/clone_1lb="GP0D"					
	/tissue_type="Immature pod walls"					
	/dev_stage="Immature pods, ranging in age from 15 to 30 days after pollination"					
	/note="Vector: plasmid SK-; Site,1: EcoRI; Site,2: XhoI; Immature pods, ranging in age from 15 to 30 days after pollination, were collected from greenhouse-grown plants. At harvest, seeds were removed from pods and isolated pod walls were collected and immediately frozen in liquid nitrogen. Pod walls were pooled for mRNA extraction. cDNA was prepared from polyA-enriched RNA. The cDNA was directionally ligated into the Uniap XR vector from Stratagene and packaged using GigaPack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda Zap phage using Ex-assist helper phage and propagated in XL0R cells."					
BASE COUNT	235 a 143 c 175 g 206 t					
Query Match	86.8%; Score 484.4; DB 13; Length 759;					
Best Local Similarity	91.8%; Pred. No. 1.2e-131;					
Matches 512; Conservative	0; Mismatches 46; Indels 0; Gaps 0;					
OY	1 TGATTTCCACCAATATTTATATGATGAGATATATTCCTCAATATTTGGAACACATGC 60					
DB	40 TGAATTCTCCACCAATATTTATATGACATGATATATTCCTCAATATTTGGAATGATGTC 99					
OY	61 GCCAAATTTGGAATAGCTGCTCATATGACCGGAAATATGCCATTATATATACAAAG 120					
DB	100 GCCAAATTTGGAATAGCTGCTCATATGCTCGGATATATGCCATTATACATTAACAG 159					
OY	121 GCTTATCCGATACACAGGCTTTTGCAAAACAAAGATACAGAGTGCATATATGCTGCT 180					
DB	160 GCTTATCCGATACATGCTTCTGTAAACAAAGATACCTGCTGAGATATGCTGCTT 219					
OY	181 TAAAGCTTATCAAGACAAATGAAGAGCTGTCAGGCTTTGGAATACACCTGGGAACCTG 240					
DB	220 TAGAGCTTATCAAGACAAATGAAGAGCTGTCAGGATGAGATATGAGATACCTGGGAACCTG 279					
OY	241 AAATTTGCAAAAAGATTTTACCTCATGATGATTAAGACATTCCTTTATACACTAAATA 300					
DB	280 AAATTTGCAAAAAGATTTTACCTCATGATGATTAAGACATTCCTTTATACACTAAATA 339					
OY	301 TGGAGAAATCTGCATTTGSCAATACTATTAACCTTTGGCTTAATGAAGAGTCCAAAGTT 360					
DB	340 TGGAGAAATCTGCATTTGSCAATCTTAACAGGCTGGGCTTAATGAAGATCCAAATTTT 399					
OY	361 CTAGAGCTTACCTTGGAGAGCGCCCTGCAGAAATGTTTCCGTGTTAAAGAAATGTCGCTC 420					
DB	400 CTAGAGCTTACCTTGGAGAGCGCCCAAAATGTTTCCGTGTTAAAGAAATGTCGCTC 459					
OY	421 CAATCTTTTGGSCAAATGACCAAAAAGGTATATCAAGAGCATATGATGATGATC 480					
DB	460 CAATCTTTTGGSCAAATGACCAAAAAGGTATATCAAGAGCATATGATGATGATC 519					
OY	481 ACCCAATATGGGCTGGGGGATTTCTGTATACCACTATAGAGGATATGATGATTC 540					
DB	520 ACCCAATATGGGCTGGAGATGATTTCCGTATACCCCTCATATGATGATTAAGTACATTC 579					
OY	541 AGTTCAATGGGCTCCACAGTC 558					
DB	580 AGTTCAATGGGCTCCACAGTC 597					

mRNA sequence.  
 ACCESSION B1308302  
 VERSION B1308302.1 GI:14982629  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
 1 (bases 1 to 755)  
 REFERENCE Grusak M.A., Smead, D.A., Town, C.D., Van Aken, S., Utterback, T., Choj and Fraser, C.M.  
 TITLE ESTs from developing reproductive tissues of Medicago truncatula  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Michael A. Grusak  
 USDA/ARS Children's Nutrition Research Center  
 Baylor College of Medicine  
 1100 Bates Street, Houston, TX 77030-2600, USA  
 Tel: 713-798-7044  
 Fax: 713-798-7078  
 Email: mgrusak@bcm.tmc.edu  
 B395157e  
 TIGR sequence name: MTOAG927K  
 More information is available at: www.medicago.org  
 Seq primer: SKmod (CTA gaa cta gta gaa cc).  
 Location/Qualifiers  
 1..755  
 /organism="Medicago truncatula"  
 /cultivar="A17"  
 /db\_xref="taxon:3880"  
 /clone="PGPD-2p15"  
 /clone\_lib="GPD"  
 /tissue\_type="Immature pod walls"  
 /dev\_stage="Immature pods, ranging in age from 15 to 30 days after pollination"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; Immature pods, ranging in age from 15 to 30 days after pollination, were collected from greenhouse-grown plants. At harvest, seeds were removed from pods and isolated pod walls were collected and immediately frozen in liquid nitrogen. Pod walls were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Uniapack III vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."  
 BASE COUNT 231 a 138 c 178 g 208 t  
 ORIGIN  
 Query Match 86.5%; Score 482.8; DB 13; Length 755;  
 Best Local Similarity 91.6%; Pred. No. 3.5e-131;  
 Matches 511; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
 Oy 1 TCGTTGTACCACTATTTTATGACGATATATTCCTCAATTTGGAACGACGTGC 60  
 Db 129 TCGTTGTACCACTATTTTATGACGATATATTCCTCAATTTGGAACGACGTGC 188  
 Oy 61 GCCAATTTGGAATAGCTGCTCATTTGCAATTTGCAATTTGCCATTAATATTCACAG 120  
 Db 189 GCCAATTTGGAATAGCTGCTCATTTGCAATTTGCAATTTGCCATTAATATTCACAG 248  
 Oy 121 GCTTATCCGATGACGTGGTCTTGGCAAAACAAAGATACAGCTGACATTTAGCTGCTT 180  
 Db 249 GCTTATCCGATGACGTGGTCTTGGCAAAACAAAGATACAGCTGACATTTAGCTGCTT 308  
 Oy 181 TAGAGCTTATCAGACATGAGAGAGCTGTCAAGCTTATGAAATTCACCTGGAACTG 240  
 Db 309 TAGAGCTTATCAGACATGAGAGAGCTGTCAAGCTTATGAAATTCACCTGGAACTG 368  
 Oy 241 AATGTCAAAAAGATTTTATGCTCATGATTAAGACATTCATTTATACCTAATAATA 300  
 |||||||||||||||| |||||||||||||||| |||||||||||||||| ||||||||

Db 369 AATGTCAAAAAGATTTTATGCTCATGATTAAGACATTCATTTATACCTAATAATA 428  
 Oy 301 TCGAAGAAATCTGCATTTGCAATGATATGACATTTGCGCTTAATTTGAAGAGTCCAAAGTTT 360  
 Db 429 TCGAAGAAATCTGCATTTGCAATGATATGACATTTGCGCTTAATTTGAAGAGTCCAAAGTTT 488  
 Oy 361 CTAGAGCTTACCTTTGAGAGACGCCCTCAAAATGTTTCCGTGTAAGAAAGATGTCCTG 420  
 Db 489 CTAGAGCTTACCTTTGAGAGACGCCCTCAAAATGTTTCCGTGTAAGAAAGATGTTTGGC 548  
 Oy 421 CAATCTTTTGGGCAATGACCAAAAGATTTGCAATTTGCAAGGACCATGATGATGATCAT 480  
 Db 549 CAATCTTTTGGGCAATGACCAAAAGATTTGCAATTTGCAAGGACCATGATGATGATCAT 608  
 Oy 481 ACCACATGCGCGGTGGGCGTATCTCTGTAATCATCATATGATGCTATGATGATTC 540  
 Db 609 ACCACATGCGCGGTGGGCGTATCTCTGTAATCATCATATGATGCTATGATGATTC 668  
 Oy 541 AGTTCATGCGCGCACGTCG 558  
 Db 669 AGTTCATGCGCGCACGTCG 686  
 RESULT 5  
 AM203249  
 LOCUS  
 DEFINITION 573 bp mRNA linear EST 02-DEC-2001  
 sf272911.y1 Gm-cl028 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl028-1245 5' similar to SW:MTNR\_ARATH 080585 PROBABLE  
 METHYLENEBIS(2-HYDROXYPROPYL) CARBONATE REDUCTASE ;, mRNA sequence.  
 ACCESSION AM203249  
 VERSION AM203249.1 GI:6501890  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 573)  
 REFERENCE Shoemaker, R., Kaim, P., Vodkin, L., Eppelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, D., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ratter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1810  
 Fax: 314 286 1800  
 Email: est@wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cdu@resgen.com  
 Insert length: 1503  
 Std Error: 0.00  
 High quality sequence stop: 405.  
 Location/Qualifiers  
 1..573  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl028-1245"  
 /clone\_lib="Gm-cl028"  
 /tissue\_type="roots of 'Superpod' plants"  
 /lab\_host="DH10b"  
 /note="Vector: pBluescript II Xr; Site\_1: EcoRI; Site\_2: XhoI; The mRNA was isolated from roots of Glycine max 'Superpod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. Stratagene's cDNA synthesis kit (catalog number 200401) was used to



```

QY 75 ACGTGCCTATTGACCTGGAATTAATGCCCATTAATAATACAGGCGCTTATCCGCAAG 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ACGAGCCCTATTGACCTGGAATTAATGCCCATTAATAATACAGGCGCTTATCCGCAAG 60
QY 135 ACTGGGTTTTCACAAACAAAGATACAGCTGACATTAATGCTGCTTTAGAGCCTTACAG 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 ACTGGGTTTTCACAAACAAAGATACAGCTGACATTAATGCTGCTTTAGAGCCTTACAG 120
QY 195 GACATGAGAAAGCTGTCAAGGCTTATGAAATTCACCTGGAACTGAAATGTCAAAAG 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GACAAAGCAAGACCTGTCAAGGCTTATGAAATTCACCTGGAACTGAAATGTCAAAAG 180
QY 255 ATTATGACCTGGAATTAAGACATTCGATTCATTAACCAATTAATGGAATTCCTCA 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 ATTATGACCTGGAATTAAGACATTCGATTCATTAACCAATTAATGGAATTCCTCA 240
QY 315 TTGGCAATACATAAGAACCTTGCCCTAATTGAGAGTCACAAAGTTCTAGTCTTACT 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 TTAGGGAATACATAAGAACCTTGCCCTAATTGAGAGTCACAAAGTTCTAGTCTTACT 300
QY 375 TTGAGAGCCCTGCAAAATGTTTCGCTGTTAAAGAGATGTCGTCGAATCTTTGGGCA 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TTGAGAGCCCTGCAAAATGTTTCGCTGTTAAAGAGATGTCGTCGAATCTTTGGGCA 360
QY 435 AATCGACCAAAAGCTACATATCAAGAGACCATAGGATGATCAATACCATAGGCGCT 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AATCGACCAAAAGCTACATATCAAGAGACCATAGGATGATCAATACCATAGGCGCT 420
QY 495 TGGGGTGATTCCTGTATCATCATATGATGATGATGATGATGATGATGATGATGATG 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 TGGGGTGATTCCTGTATCATCATATGATGATGATGATGATGATGATGATGATGATG 480
QY 555 CGTG 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 CGTG 484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
LOCUS BG405762 468 bp mRNA linear EST 28-NOV-2001
DEFINITION sac37f08.y1 Gm-cl051 glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl051-4503 5' similar to SW:MTFR_ARAH 080585 PROBABLE
METHYLMETHYLETETRAHYDROFOLATE REDUCTASE ; mRNA sequence.
BG405762
ACCESSION BG405762.1 GI:1312111
VERSION BG405762.1
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
            Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
            Spermatophyta: Magnoliophyta: eudicotyledons: Core eudicots:
            Rosidae: eurosids I: Fabales: Fabaceae: Papilionoideae: Phaseoleae:
            Glycine.
            1 (bases 1 to 468)
REFERENCE 1
  Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
  ,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,D., Beck,C.,
  Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
  ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
  ,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
  ,R., Waterston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished (1999)
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@wustl.edu
  This clone is available through: Resgen, Invitrogen Corp. 2130
  South Memorial Parkway Huntsville, AL 35801 For further information
  call: (800)-533-4363 or contact via email: cou@resgen.com
  High quality sequence stop: 338.

```

```

FEATURES
    source          location/Qualifiers
                    1..468
                    /organism="Glycine max"
                    /db_xref="taxon:3847"
                    /clone="GENOME SYSTEMS CLONE ID: Gm-cl051-4503"
                    /clone_id="Gm-cl051"
                    /tissue_type="floral meristematic mRNA"
                    /lab_host="DH10B"
                    /note="Vector: pBluescript II SK(+), Site_1: EcoRI, Site_2:
                    XhoI. The cDNA library was constructed from floral
                    meristematic mRNA provided by Dr. Halina Knap of Clemson
                    University. Complementary DNA was synthesized from mRNA
                    using a primer consisting of a poly(dT) sequence with a
                    XhoI restriction site. EcoRI adapters were ligated to the
                    blunt-ended cDNA fragments followed by XhoI digestion. The
                    cDNA fragments were directionally cloned into the
                    EcoRI-XhoI restriction site of the pBluescript vector. The
                    ligated cDNA fragments were transformed into DH10B host
                    cells (GibcoBRL). This library was constructed in the
                    laboratory of Dr. Randy Shoemaker."

BASE COUNT      146 a      87 c      96 g      139 t

ORIGIN
Query Match      77.0%; Score 429.4; DB 12; Length 468;
Best local similarity 95.5%; Pred. No. 1.5e-115;
Matches 442; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TGATGTCAACCAATTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 TGATGTCACTAGTATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 64
QY 61 GCCAATTTGGAATTAACGATGCTGATGATGATGATGATGATGATGATGATGATGATG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 GCCAATTTGGAATTAACGATGCTGATGATGATGATGATGATGATGATGATGATGATG 124
QY 121 GCTTTATCCGATGACTGGGTTTGCACAAAGATACAGCTGACATTAATGCTGCTT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 GCTTTATCCGATGACTGGGTTTGCACAAAGATACAGCTGACATTAATGCTGCTT 184
QY 181 TAGAGCCTATCAAGACAAAGAGCTGTCAAGGCTTATGGAATTCACCTGGGAATG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 TAGAGCCTATCAAGACAAAGAGCTGTCAAGGCTTATGGAATTCACCTGGGAATG 244
QY 241 AATGTGCAAAAGATTTTACCTGATGATGATGATGATGATGATGATGATGATGATG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 AATGTGCAAAAGATTTTACCTGATGATGATGATGATGATGATGATGATGATGATG 304
QY 301 TGGAGAAATCTGATGCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 305 TGGAGAAATCTGATGCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 364
QY 361 CTGAGTCTTACCTTGGAGAGCGCCCTGCAAAATGTTTCCGCTTTAAAGAGATGTCG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 CTGAGTCTTACCTTGGAGAGCGCCCTGCAAAATGTTTCCGCTTTAAAGAGATGTCG 424
QY 421 CAATCTTTTGGGCAATGACCAAAAGCTACATATCAAGAC 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 CAATCTTTTGGGCAATGACCAAAAGCTACATATTAAGAC 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
LOCUS BQ404086 668 bp mRNA linear EST 22-MAY-2002
DEFINITION GA_Ed0066A09f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ed0066A09f, mRNA sequence.
BQ404086
ACCESSION BQ404086.1 GI:21091773
VERSION BQ404086.1
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum
            Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
            Spermatophyta: Magnoliophyta: eudicotyledons: Core eudicots:
            Rosidae: eurosids II: Malvales: Malvaceae: Gossypium.

```

REFERENCE	1 (bases 1 to 668)
AUTHORS	Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
TITLE	An integrated analysis of the genetics, development, and evolution of the cotton fiber
JOURNAL	Unpublished (2000)
COMMENT	Contact: Wing RA Clemson University Genomes Institute Clemson University 100 Jordan Hall Clemson, SC 29634, USA Tel: 864 656 7286 Fax: 864 656 4293 Email: rwing@clemson.edu
FEATURES	total High Quality bases = 504 Seq primer: TAAATGCTCACTAAGG High quality sequence start: 14 High quality sequence stop: 653. Location/Qualifiers
SOURCE	1..668 /organism="Gossypium arboreum" /strain="AXA" /cultivar="8400" /db_xref="taxon:29729" /clone="GA_Ed0066A09f" /clone_1fb="Gossypium arboreum 7-10 dpa fiber library" /hisue_type="Fibers isolated from bolls harvested 7-10 dpa" /lab_host="E. coli" /note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	197 a 131 c 143 g 197 t
ORIGIN	
Query Match	75.64; Score 422; DB 14; Length 668;
Best Local Similarity	84.88; Pred. No. 2.6e-113;
Matches 473; Conservative	0; Mismatches 83; Indels 0; Gaps 0;
OY	1
DB	55
OY	61
DB	115
OY	121
DB	175
OY	181
DB	235
OY	241
DB	295
OY	301
DB	355
OY	361
DB	415
OY	421
DB	475
OY	481
DB	535
OY	541

Db	595	AGTCATGAGCGCTCGAG	612
RESULT 9	BP269445	589 bp	MRNA linear
LOCUS	GA_Eb0004L07f	Gossypium arboreum 7-10 dpa fiber library	Gossypium
DEFINITION	arboresum cDNA clone GA_Eb0004L07f,	core	eudicotis;
ACCESSION	BP269445		
VERSION	BP269445.1	GI:11200440	
KEYWORDS			
SOURCE			
ORGANISM	Gossypium arboreum.		
REFERENCE	Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry		
AUTHORS	'D., Wood,T.C., Leslie,A. and Wilkins,T.A.		
TITLE	An integrated analysis of the genetics, development, and evolution		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Wing RA		
ORIGIN	Clemson University Genomics Institute		
	100 Jordan Hall, Clemson, SC 29634, USA		
	Tel: 864 656 7288		
	Fax: 864 656 4293		
	Email: rwing@clemson.edu		
	Seq primer: TAATACGACTCACTATAGG		
	High quality sequence stop: 585.		
FEATURES			
Source	1..589		
	/organism="Gossypium arboreum"		
	/strain="AKA"		
	/cultivar="8400"		
	/db_xref="taxon:29729"		
	/clone="GA_EB0004L07f"		
	/cclone_lib="Gossypium arboreum 7-10 dpa fiber library"		
	/rclone_type="fibers isolated from bolls harvested 7-10		
	dpa		
	/lab_host="E. coli"		
	/note="Vector: pBK-CMV; Site_1: EcoRI, Site_2: XhoI"		
BASE COUNT	170 a 112 c 125 g 180 t	2 others	
ORIGIN			
Query Match	71.8%;	Score 400.4;	DB 12; Length 589;
Best Local Similarity	85.3%;	Pred. No. 5.7e-107;	
Matches 446; Conservative	0;	Mismatches 77;	Indels 0; Gaps 0.
1	TGATGTGACCAATTAATTTATATATACGAGATATATTCCTCAATTTGGAACGACTGTC	60	
2			
3			
4			
5			
6	TGATTTTATCTACACGTTTATATATATACGATATATTTCTCAATTTGTCATGATGCTC	125	
7			
8			
9			
10			
11			
12	GCCAAATTTGAAATTAATTTATGTTCTCTATGCTGCGGATATATATCCATTAAATACCTCAAG	185	
13			
14			
15			
16	GCTTATGCCCATGACTGSGTGTGGCAAAACAAAGATATACAGCTGACATTAATGCGTCTT	180	
17			
18			
19			
20			
21	GTTTCTTGGCAATGAGCTGGCTTTGGCAAAATACATAGACAGCTGAGATTAATCTGCTCT	245	
22			
23			
24			
25			
26	TGAGAGCTTACACAGCAATAGACACTGCTCAAGCTTATGAGAAATTCACCTGGAGCTG	240	
27			
28			
29			
30			
31	TAGAGCTTACACAGCAACACGAAACACTCTCTCAAGCCTATAGAGTTCAACCTTGGACAG	305	
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			

Db 366 TGGAGAAATCGATTAGCTATCTTAATGAATCTTGATTTATGAAGTCCAAATTT 425  
 Oy 361 CTAGTCCCTTACCTTGGAGAGCCCTGCAAAATGTTTTCGGTTTAAAGAGATGCCGTC 420  
 Db 426 CTAGTCCCTTACCTTGGAGAGCCCTGCAAAATGTTTTCGGTTTAAAGAGATGCCGTC 485  
 Oy 421 CAATCTTTGGGCAATCGACCAAAAGCTATCATCAGAGCATAGATGGATCAAT 480  
 Db 486 CAATCTTTGGTCTAACCCCTCTAGAGATTAACAACGAGCATATGGTTGTGATCACT 545  
 Oy 481 ACCCATATGGGCGTTGGGCGTATCTCTTAATCATCATATATG 523  
 Db 546 ACCCATATGGGCGATGCTGATCTCTTAATCATCATATATG 588  
 RESULT 10  
 BE204232 593 bp mRNA linear EST 05-SEP-2000  
 LOCUS BE204232 593 bp mRNA linear EST 05-SEP-2000  
 DEFINITION EST396908 KV0 Medicago truncatula cDNA clone pkV0-14L8, mRNA  
 sequence.  
 ACCESSION BE204232  
 VERSION BE204232.1 GI:8747517  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 1 (bases 1 to 593)  
 Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,  
 Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and  
 Fraser, C.M.  
 ESTs from uninoculated seedling roots of Medicago truncatula  
 Unpublished (1999)  
 CONTACT: Vandenbosch K  
 Department of Plant Biology  
 University of Minnesota  
 220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA  
 Tel: 612 624 2755  
 Fax: 612 625 1738  
 Email: kvandenbosch.umn.edu  
 Texas A&M University name: T264004e  
 TIGR sequence name: MGC64647  
 More information is available at:  
 http://chrysis.tamu.edu/medicago  
 Seq primer: SKmod (CTA GAA CTA gtg gat CC).  
 Location/Qualifiers  
 1..593  
 /organism="Medicago truncatula"  
 /cultivar="genotype A17"  
 /db\_xref="taxon:3880"  
 /clone="pkV0-14L8"  
 /clone\_lib="KV0"  
 /tissue\_type="Seedling roots"  
 /dev\_stage="Immediately prior to inoculation with  
 Sinorhizobium meliloti (0 hour)"  
 /lab\_host="E.coli strain XL0LR"  
 /note="Vector: pBluescript SK-; Site: 1: EcoRI; Site: 2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA  
 was directionally ligated into the Unizap XR vector from  
 StrataGene and packaged using Gigapack III Gold packaging  
 extracts. Plasmids containing cDNA inserts were excised  
 from the recombinant lambda-Zap phage using Ex-assist  
 helper phage and propagated in XL0LR cells."

BASE COUNT 189 a 113 c 137 g 154 t  
 ORIGIN  
 Query Match 71.5%; Score 399.2; DB 10; Length 593;  
 Best local similarity 90.8%; Pred. No. 1.3e-106;  
 Matches 425; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
 Oy 91 CTGGAATATGCCATTATATATATACAAAGGCTTATCCGATACACGGGTTTGCAAAA 150

Db 1 CTGGAATATGCCATTATATATATACAAAGGCTTATCCGATACACGGGTTTGCAAAA 60  
 Oy 151 CAAGATACACCTGACATATATGCTGCTTTAGAGCCCTATCAAGCAATGAGACCTG 210  
 Db 61 CAAGATATGCTGCTATATATATGCTGCTGCTGAGCCCTATCAAGCAATGAGACCTG 120  
 Oy 211 TCAAGGCTTATGAAATTCACCTGGGAGATGAAATGTCAGAAAAAGATTTTACCTGAA 270  
 Db 121 TCAAGGCTTATGAAATTCACCTGGGAGATGAAATGTCAGAAAAAGATTTTACCTGAA 180  
 Oy 271 TTAAGCATTTGATCTTATATACACTAAATATGAGAAATATGCTATGCAATATGTA 330  
 Db 181 TTAAGCATTTGATCTTATATACAGCTTAAATATGAGAAATATGCTATGCAATATGTA 240  
 Oy 331 ACCTGGGCTTAAATGAGAGTCCAAAGTTCTAGGCTCTTACCTTGGAGAGCCCTGCA 390  
 Db 241 GCCTGGGCTTAAATGAGAGTCCAAAGTTCTAGGCTCTTACCTTGGAGAGCCCTGCA 300  
 Oy 391 ATGTTTCCGTTTAAAGAGATGCTGCTGCTTATGAGCAATGCTTGGCAATGCAAAAAGCT 450  
 Db 301 ATGTTTCCGTTGAAAGAGATGCTGCTGCTTATGAGCAATGCTTGGCAATGCAAAAAGCT 360  
 Oy 451 ACATATCAAGACCATATGAGATGATGATACCATACCATGAGGCGATGATTCCTGTA 510  
 Db 361 ACATATCAAGACCATATGAGATGATGATACCATACCATGAGGCGATGATTCCTGTA 420  
 Oy 511 ATCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 558  
 Db 421 ACCGTCATATGCTGATTAAGTACTATCACTATCACTATCACTATCACTATCACTATCA 468  
 RESULT 11  
 BE433609 558 bp mRNA linear EST 18-MAY-2001  
 LOCUS BE433609 558 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST400138 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
 clone cBE43602, mRNA sequence.  
 ACCESSION BE433609  
 VERSION BE433609.1 GI:9431452  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 1 (bases 1 to 558)  
 Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,  
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Rouning, C.M.,  
 Nierman, M., Fraser, C.M., Martin, G.B., Giovannoni, J.V. and Tanksley  
 S.D.  
 Generation of ESTs from tomato fruit tissue, breaker stage  
 Unpublished (2000)  
 CONTACT: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: http://www.genome.clemson.edu/orders/index.html  
 5 prime sequence.  
 Location/Qualifiers  
 1..558  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cBE43602"  
 /clone\_lib="tomato breaker fruit, TIGR"  
 /tissue\_type="Pericarp"  
 /dev\_stage="Breaker"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescriptSKmChadp1; Site: 1: EcoRI;  
 Site: 2: XhoI; Fruit were harvested at the breaker stage  
 (first sign of lycopene accumulation on the blossom end of  
 the fruit). Fruit were cut in half and the seeds and











Qy	241	QLEKINTNGELTITINOSP	PAYNGEKEDSDTPTVMGSPGGYVYQKAAVEFFCSKEKIDLADYK	300
Dy	427	QIGKINSNGELTITINOSP	ANNAKSDPAAIDMGPGGYVYQAAITFEFCSKDKLIDLAYKS	486
Qy	301	KRTSTLTVAVNGDSKSN	QNGKIDBPAIDMGPGGYVYQAAITFEFCSKDKLIDLAYKS	366
Dy	487	KAFSTITVAVNGDSKSN	QNGKIDBPAIDMGPGGYVYQAAITFEFCSKDKLIDLAYKS	546
Qy	361	MARGWASLPEDDEASKRL	ELVEEGSHPLVLDVNDINDGDLFAVFA	406
Dy	547	WRSWANLPEDDEASKRL	ELVEEGSHPLVLDVNDINDGDLFAVFA	592

```

RESULT 2
US-09-931-795-4
: Sequence 4, Application US/09931795
: Publication No. US20020198211A1
: GENERAL INFORMATION:
: APPLICANT: ROZEN, Rima
: TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
: FILE REFERENCE: 0484/005003
: CURRENT APPLICATION NUMBER: US/09/931.795
: CURRENT FILING DATE: 2001-08-16
: PRIOR APPLICATION NUMBER: US 09/592,595
: PRIOR FILING DATE: 2000-06-12
: PRIOR APPLICATION NUMBER: US 09/258,928
: PRIOR FILING DATE: 1999-03-01
: PRIOR APPLICATION NUMBER: US 08/738,000
: PRIOR FILING DATE: 1997-02-12
: PRIOR APPLICATION NUMBER: PCT/K95/00314
: PRIOR FILING DATE: 1995-05-25
: PRIOR APPLICATION NUMBER: GB 9410620.0
: PRIOR FILING DATE: 1994-05-26
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 656
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-931-795-4

```

[illegible]

DB 580 FFMFKMDEAFALMIEKMGKLYEESSPSRITLYIHDNYEIVLWVNDPFLDNCMQVED 638

RESULT 3  
US-09-728-910-4  
Sequence 4, Application US/09728910  
Patent No. US20010025030A1  
GENERAL INFORMATION:  
APPLICANT: Rozen, Rima  
APPLICANT: Sekhon, Jaspreet  
TITLE OF INVENTION: REDCA FOR HUMAN METHYLENETETRAHYDROFOLATE  
FILE REFERENCE: 04844/006001  
CURRENT FILING DATE: 2000-12-01  
CURRENT APPLICATION NUMBER: US/09/728,910  
PRIOR FILING DATE: 1999-03-01  
PRIOR APPLICATION NUMBER: US 09/258,928  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 656  
TYPE: FRT  
ORGANISM: Homo sapiens  
US-09-728-910-4

```

QY      1 ARIYVOLFETYDIPFLIFENDRCRIGITCPIVPGIMPINNKGFIRMTFOFKTIPADIM 60
    Db   222 ADRIITQLEFFADYFPFRFACCTDMGITCPIVPGIIGHSURDYAKLKEVPEIK 281
QY      61 AALEPKINENKAAVAGIHLDTCEKKILANG-ITKLHLYTLNNKSALATIMMLGIREE 119
    Db   282 DYIEPKKNDAARNGYIELVASCOELLASGLAPGRLHETYLINENMATTVEVKRLGMTE 341
QY     120 SKYSRSLEWRPMPANFYRYAKEVRPFIFMARNSKSYISRTIGMDQYPGRMGSCSPSGAL 179
    Db   342 DP-RRLFWALSHPRRREEDVRPIFMASRKSYIYRDNDDEDFRGWSSSFGEL 400
QY     180 SDVO--FMRPRADKKLVEMEAAYPLKSVSDIYEAFYLCGLKLSN-----PWSELDG 230
    Db   401 KDYLEEYLNKSKREBELTKMMGEELISEAVFEYFVLAYSGBEPNRKHKVTCLPMD-SP 459
QY     231 LDPETKIINEOLKIRINTGSFLTINSOPAVNEKSDSPYTGNGSGPGGYXOKRAYEFCCK 290
    Db   460 LAERYSLLKRELRLRVNRGGILTITNSOPINORKBSPDIYMGPSGGGYEQAKYLEFTFSR 519
QY     291 EKLDALYDKCD-RTSLTYMAVVKDSMSNVGQTGVNAVYMGVFANEIIIOPTIDVPS 349
    Db   520 ETWEALDGVLKKEYELRVNHLYVNKGENTTAAPDLPAHYVMWGJFGREITTIQPVDPVS 579
QY     350 FYNWKDEAFETINSRWASLYPEDEASRKLYEEVGSHLEVSLVANDY-INGDLFAVAD 407
    Db   580 FMKWDEAFALMIERNKGLYEESSPERTIQLYHDNYPLVNLVNDPFLDNCIMQVVED 638

RESULT 4
US-09-931-795-2
Sequence 2, Application US/09391795
Publication No. US2002019821A1
GENERAL INFORMATION:
APPLICANT: KOZEN, Rama
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROLATE
FILE REFERENCE: REDUCTASE AND USES THEREOF
CURRENT APPLICATION NUMBER: US/09/931,795
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: US 09/592,595
PRIOR FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: US 09/258,928
PRIOR FILING DATE: 1999-03-01

```

PRIOR APPLICATION NUMBER: US 08/738,000  
PRIOR FILING DATE: 1997-02-12  
PRIOR APPLICATION NUMBER: PCT/CA95/00314  
PRIOR FILING DATE: 1995-05-25  
PRIOR APPLICATION NUMBER: GB 9410620.0  
PRIOR FILING DATE: 1994-05-26  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 660  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-09-931-795-2

Query Match 44.8%; Score 979; DB 9; Length 660;  
Best Local Similarity 44.9%; Pred. No. 7.2e-85;  
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;

QY 1 ARVLTQLEFYDIDFLKFNDCRQIGTCPIVPGIMPINNYKGFIRMTGCKTKIPADIM 60  
DB ADITITQLEFADTFPRFVYKACCTDMGICPIVPGIFPIQGYHSLRQLVLSKLEVPQEK 285  
QY 61 AALEPINDNEAKAYGILHGTGEMCKKILAHG-IKTLHLTYLNNKESALAILNNLGLIEE 119  
DB DVEPIKNDNAALRNIGIELAVSLCOELASGLVGLHFFLYLNKEMATTEYLKRLGMTE 345  
QY 120 SKVRSLLPWRPAPNVRKEDVRPIFWANRPKSYISRTIGMDQYPHGRMGDSNPSYCAL 179  
DB 346 DP-RRPLPMAISAPKREEDVRPIFWASRPKSYITRQEWDEPPNGMSSPARGEL 404  
QY 180 SDVQ-FMRBRADKLYEMAVPLKSVEDIYERRLYCLGLKLSN-----PMSLDS 230  
DB 405 KDYLFLTKSKSKEELLMKMBEELTSEASVEFVLTLSCEPNRNHAKYTCLPMND-EP 463  
QY 231 LQPETKILNOLKIKNTKGLTINSOPAVNGEKSDPTVWGSGGYVYOKAYVEFFCSK 290  
DB 464 LAETSLKEELLRVNRQGLITINSQPNNGKPSDPIVWGSGGYVYOKAYLEFFTSR 523  
QY 291 EKLDALVDKCD-RTSLTYMAVNNKDSKSNVCGTDVAATVWGVFPAKELIOTPIVDPVS 349  
DB 524 ETAEALQVLAKYELRVNHLVNVGENTNAPELQPMAYTWGIFPGRREIIOPTVDPVS 583  
QY 350 FNVKDEAFETISRGWASLYPEDEASRKLYEVEGSHPLVSLVNDY-INGDLFAVPAD 407  
DB 584 FMEWKDEAFALMIERWKGKLYEESPSRTIYIHDNYFLVNLVNDPDLNCLMOWVED 642

## RESULT 5

US-09-728-910-2  
Sequence 2, Application US/09728910  
Patent No. US20010025030A1  
GENERAL INFORMATION:  
APPLICANT: Rozen, Rima  
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE  
TITLE OF INVENTION: REDUCTASE AND USES THEREOF  
FILE REFERENCE: 04844/006001  
CURRENT APPLICATION NUMBER: US/09/728,910  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: US 09/258,928  
PRIOR FILING DATE: 1999-03-01  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 660  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-09-728-910-2

Query Match 44.8%; Score 979; DB 10; Length 660;  
Best Local Similarity 44.9%; Pred. No. 7.2e-85;  
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;

QY 1 ARVLTQLEFYDIDFLKFNDCRQIGTCPIVPGIMPINNYKGFIRMTGCKTKIPADIM 60  
DB ADITITQLEFADTFPRFVYKACCTDMGICPIVPGIFPIQGYHSLRQLVLSKLEVPQEK 285  
QY 61 AALEPINDNEAKAYGILHGTGEMCKKILAHG-IKTLHLTYLNNKESALAILNNLGLIEE 119  
DB DVEPIKNDNAALRNIGIELAVSLCOELASGLVGLHFFLYLNKEMATTEYLKRLGMTE 345  
QY 120 SKVRSLLPWRPAPNVRKEDVRPIFWANRPKSYISRTIGMDQYPHGRMGDSNPSYCAL 179  
DB 346 DP-RRPLPMAISAPKREEDVRPIFWASRPKSYITRQEWDEPPNGMSSPARGEL 404  
QY 180 SDVQ-FMRBRADKLYEMAVPLKSVEDIYERRLYCLGLKLSN-----PMSLDS 230  
DB 405 KDYLFLTKSKSKEELLMKMBEELTSEASVEFVLTLSCEPNRNHAKYTCLPMND-EP 463  
QY 231 LQPETKILNOLKIKNTKGLTINSOPAVNGEKSDPTVWGSGGYVYOKAYVEFFCSK 290  
DB 464 LAETSLKEELLRVNRQGLITINSQPNNGKPSDPIVWGSGGYVYOKAYLEFFTSR 523  
QY 291 EKLDALVDKCD-RTSLTYMAVNNKDSKSNVCGTDVAATVWGVFPAKELIOTPIVDPVS 349  
DB 524 ETAEALQVLAKYELRVNHLVNVGENTNAPELQPMAYTWGIFPGRREIIOPTVDPVS 583  
QY 350 FNVKDEAFETISRGWASLYPEDEASRKLYEVEGSHPLVSLVNDY-INGDLFAVPAD 407  
DB 584 FMEWKDEAFALMIERWKGKLYEESPSRTIYIHDNYFLVNLVNDPDLNCLMOWVED 642

## RESULT 6

US-10-113-852A-2  
Sequence 2, Application US/10113852A  
Publication No. US20020192784A1  
GENERAL INFORMATION:  
APPLICANT: Appling, Dean R.  
APPLICANT: Hanson, Andrew D.  
APPLICANT: Raymond, Rhonda R.  
APPLICANT: Roje, Sanja  
TITLE OF INVENTION: Biosynthesis of S-Adenosyl Methionine in a Recombinant Yeast S.  
FILE REFERENCE: 119927-1033  
CURRENT APPLICATION NUMBER: US/10/113,852A  
CURRENT FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: 60/280,333  
PRIOR FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 599  
TYPE: PRF  
ORGANISM: Saccharomyces cerevisiae  
US-10-113-852A-2

Query Match 35.6%; Score 778; DB 9; Length 599;  
Best Local Similarity 40.9%; Pred. No. 8.3e-66;  
Matches 168; Conservative 76; Mismatches 147; Indels 20; Gaps 11;

QY 4 IVTQLEFYDIDFLKFNDCRQIGTCPIVPGIMPINNYKGFIRMTGCKTKIPADIM 63  
DB IITQLEFYDIDFLKFNDCRQIGTCPIVPGIMPINNYKGFIRMTGCKTKIPADIM 244  
QY 64 EPIKNDNEAKAYGILHGTGEMCKKILAHG-IKTLHLTYLNNKESALAILNNLGLIEE 121  
DB 245 DVEPIKNDNAALRNIGIELAVSLCOELASGLVGLHFFLYLNKEMATTEYLKRLGMTE 304  
QY 122 VSRSLPWRPAPNVRKEDVRPIFWANRPKSYISRTIGW--DQYPHGRMGDSNPS 175  
DB 305 FNAHPLAVLPMRKSLLPKRNEVEPIFWKRBPYVARTSQMVADEPPNGRFGDSSPA 364  
QY 176 YGAL----SDVQFMRBRADKLYEMAVPLKSVEDIYERRLYCLGLKLSNPMSELDG 231  
DB 365 FGDDLCLGSD--ILROSA--NCKLELWSTP-TSINDVAFVINTINGKCLKLPMSDLP-1 418  
QY 232 QPETKILNOLKIKNTKGLTINSOPAVNGEKSDPTVWGSGGYVYOKAYVEFFCSK 291

```
Db 419 NDEINPKAHLELNQSIITINSQPVNGISNKGIMGPKGQYQKQYLEFM.LPKT 478
Qy 292 KLDLAVKCKDRSLTYMAMNKGSKSN-VQOTVNAVTVGVPAPKEIIPTIVDVVSF 350
Db 479 LKPLITDPLKNNNEPLTFALDSQGLLSNHPNKSNNVMTWGIPEGREILOPTIVEXISF 538
Qy 351 NWKDEAFETWSRGMSLYPEDEA-SRKLVEVGGSHPLVSLVDNDYINGD 400
Db 539 LAKKEEYHILNEMKLNMMNKYDKPHGACFIQSLIDYCLVNIVDNDYISPD 589
```

## RESULT 7

```
US-09-734-017A-62
; Sequence 62, Application US/09734017A
; Patent No. US20020142422A1
; GENERAL INFORMATION:
; APPLICANT: Lersch, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reinold, Andreas
; APPLICANT: Cirus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides and
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/171,100
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect
; SEQ ID NO 62
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-017A-62
```

```
Query Match 33.9%; Score 742; DB 10; Length 191;
Best Local Similarity 79.9%; Pred. No. 4,1e-63;
Matches 135; Conservative 14; Mismatches 20; Indels 0; Gaps 0;
```

```
Qy 3 VIVTQLFYDIDFLKFNDCRQIGTCPIVPGIMPINNYKGFIRMTGFCFKIPADIMAA 62
Db 3 VIVTQLFYDIDFLKFNDCRQIGTCPIVPGIMPINNYKGFIRMTGFCFKIPADIMAA 62
Qy 63 LEPTKDNKAEVAVGIGHGTEMCKKILANGIKTLHLTYLNLNKSALIMLMGLIESKY 122
Db 63 LEPTKDNKAEVAVGIGHGTEMCKKILANGIKTLHLTYLNLNKSALIMLMGLIESKY 122
Qy 123 SRSILPMRRPANVFRKEDVAPITFMANRPKSYISRTIGMDQYPHRGWDS 171
Db 123 SRSILPMRRPANVFRKEDVAPITFMANRPKSYISRTIGMDQYPHRGWDS 171
```

## RESULT 8

```
US-09-734-017A-64
; Sequence 64, Application US/09734017A
; Patent No. US20020142422A1
; GENERAL INFORMATION:
; APPLICANT: Lersch, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reinold, Andreas
; APPLICANT: Cirus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
```

```
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides an
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/171,100
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect
; SEQ ID NO 64
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-017A-64
```

```
Query Match 9.1%; Score 199; DB 10; Length 136;
Best Local Similarity 90.2%; Pred. No. 1.2e-11;
Matches 37; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 3 VIVTQLFYDIDFLKFNDCRQIGTCPIVPGIMPINNYK 43
Db 87 VIVTQLFYDIDFLKFNDCRQIGTCPIVPGIMPINNYK 127
```

## RESULT 9

```
US-09-738-626-5879
; Sequence 5879, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: MAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5879
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5879
```

```
Query Match 7.2%; Score 158; DB 9; Length 326;
Best Local Similarity 27.7%; Pred. No. 3.3e-07;
Matches 36; Conservative 28; Mismatches 46; Indels 20; Gaps 4;
```

```
Qy 1 ARVIVTQLFYDIDFLKFNDCRQIGTCPIVPGIMPINNYKGFIRMTGFC 51
Db 189 AYSITQMFEDVDYLRCLDRVLAADPIHGAK-----PIIPGIMPITSRSVRRQVEIS 242
Qy 52 KTKIPADIMAA-EPKDNKAEVAVGIGHGTEMCKKILANGIKTLHLTYLNLNKS 106
Db 243 GAQIPSOLESILVRAANGNEANKDEIRKVGIEYSTNAERLIAEGADLHFTLNFTRA 302
```



TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium falciparum  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-153-273-12

Query Match 4.5%; Score 97.5; DB 9; Length 2710;  
Best Local Similarity 19.8%; Pred. No. 4.3;  
Matches 88; Conservative 59; Mismatches 142; Indels 155; Gaps 27;

QY 56 PADIMAAAL-EPKIDNEAVKAVGILHGTMM---CKIILAHGKILHLTLTMMEKSAALAIL 111  
DB 182 PSQCTWLARSFADIDIVRGDLVGNPQELKQOLENNILKTL----- 226  
QY 112 MNLGLIEESKVSRLPWRPRANVFRKEDVRPIFW-ANRPKSYISRTIGMDQYPHGMGD 170  
DB 227 --FGKIYEKINGAARVAGNDPEEFKLRD---WWTANR-----ETV-WKAILTCNMGN 273  
QY 171 S-----CN-----PSYCALSDY--QFMRPRARDKKLVEEAVL----- 202  
DB 274 TYFHATCNRGERTKGYCRCDNDQVPTV--FDYVPOYLK-----WFEWAEDFCRRKN 323  
QY 203 KSVEDIYERFR-----LYC-----LGKLR-----SNPW----- 225  
DB 324 KRIKDVARNCRGKDEKEDKRYGSRNGYDCEKTKRAIGLRGKQICISCLYACNRYVDWIN 383  
QY 226 SELDGLQPEKTIINOLEKINTKGLTITSQPAVNGEKSD---SPTVGMGSPGGYVYOKA 282  
DB 384 NQKQFQKQKKYDEETKYE-----NGASGSSRQKRDAGTTTNYDQ-----YEKK 431  
QY 283 Y-----VEFCSKEKLDAVLCCKDRT--SLTYMAVNMKDSWKSNAVQOTDYNVAV 329  
DB 432 FYDELANKSEYRTVDKFLKLSNNEICTKVKDEGGTIDFKNNNSD---SISGASGATNVE 487  
QY 330 TWGTFPAKKEITQP-----TIYDPSFVWAKDEAFELTMSKMASLY---PEDASRKLV 379  
DB 488 SQGFYFSKVCPCPYCGVAKVNNNGSSNMEMEKNNKCKSG--KLYPEKPKDEGTTTII 545  
QY 380 EEVGSHPFLVSLVDN---DYINGD 400  
DB 546 LKSGKHGHDIEKLKMFCEKKNKD 569

RESULT 13  
US-09-911-888-23  
Sequence 23, Application US/09911888  
Patent No. US20020119509A1  
GENERAL INFORMATION:  
APPLICANT: Kaitlin, Yigal  
APPLICANT: Gavriel, Victoria  
TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE  
FILE REFERENCE: 06286-062002  
CURRENT APPLICATION NUMBER: US/09/911,888  
PRIOR FILING DATE: 2001-07-23  
PRIOR APPLICATION NUMBER: US 08/965,762  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 23  
LENGTH: 347  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-911-888-23

Query Match 4.2%; Score 91; DB 10; Length 347;  
Best Local Similarity 21.3%; Pred. No. 0.87;  
Matches 76; Conservative 57; Mismatches 118; Indels 114; Gaps 22;

QY 2 RVIYTOLEFYTDIDFLKAVNDRCRIGITCPRYGIMINNYKGFIMTQCKTKPADIMA 61  
DB 22 RLIEFE--DTNRLTIVTNSR---TLPRVQDV--INQIKFYKSRNVE-DLEIDPDY 71  
QY 62 ALEPIKIDNEAVKAVGILHGTCEMKKILAHGKILHLTLTMMEKSAALAIL-MN----- 113  
DB 72 LLVDFTMNAVSYLVANV-----YQINKRYAIRNLFVMAAQGFED 109  
QY 114 ----LGILIEESKVSRLPWRPRANVFRV-----KEVRIPIFAN--RKSYSIRIGMDQ 162  
DB 110 GIDWIGAVKEVETNPDLAVTNPTPKYIQLGVKSKDDMGILFOANVFGPYEFSKTL--PQ 167  
QY 163 YPHGR---WQDS--CNPSYCALSDYQFMRPRAR--DKLVEEAVPLKSVEDI----- 208  
DB 168 LTRGKATYVWISSIMSDPKLSINDIELLKTNSYEGSKRLVD--LHLATYKDLKTKGI 225  
QY 209 -----YERFRLYC---LGKLRSPNSLGLDLOPE-----TKIT 238  
DB 226 NQYVVOPIFTSHSFSSEYLNFFTFGMLCLFYLARLLGSPMHNIDGYAANAAPVYVTRLA 285  
QY 239 NEOLEKINTK-GFLTI--NSQPAVNGEKSDSPVGMGSPGGYVYOKAYVEFCSKEKLDAI 296  
DB 286 NPNEFKQDVKGSATSRDGMPIKIQEID-----PTGMSDVFAVYIO---KKKLE-W 332  
QY 297 VDKCKDR 303  
DB 333 DEKLKQ 339

RESULT 14  
US-10-003-392-2  
Sequence 2, Application US/10003392  
Patent No. US20020170086A1  
GENERAL INFORMATION:  
APPLICANT: Allen, Stephen M.  
APPLICANT: Caiml, Perry G.  
APPLICANT: Stoop, Johan M.  
TITLE OF INVENTION: Fructan Biosynthetic Enzymes  
FILE REFERENCE: BR1463 US NA  
CURRENT APPLICATION NUMBER: US/10/003,392  
PRIOR FILING DATE: 2001-10-30  
PRIOR APPLICATION NUMBER: 60/244,273  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: 60/269,543  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 2  
LENGTH: 608  
TYPE: PRT  
ORGANISM: Dimorphotheca sinuata  
US-10-003-392-2

Query Match 4.2%; Score 91; DB 9; Length 608;  
Best Local Similarity 18.9%; Pred. No. 2;  
Matches 86; Conservative 69; Mismatches 150; Indels 150; Gaps 25;

QY 60 MALEPIKIDNEAVKAVGILHGTCEMKKILAH-----GIKTLHLTMMEKSA 107  
DB 1 MTTTKFESDLEDA-----PLINHTREPPPPPTTAGRKLLIIVYVITLL 46  
QY 108 LAILMNLGLIEESKVSRS-----LPWRPRANVFRKEDVRPIFWANRPKSYIS 155  
DB 47 ILLIYSVLFLNQONSSHSTNSKSISSDRLIWERTSFHFQPAKNF--IYDNGFLFH-- 102  
QY 156 RTIGW---DQY-PHGR-WQDSQNPYCALSDYQFMRPRARDKKLVEEAVV----- 201  
DB 103 --MGWYHLFYQINPYGPVWG--NMSWG-----HSVSKDKIMNDELVALVPTFW 147





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 12, 2003, 17:03:12 : Search time 75 Seconds

(without alignments)  
2770.721 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187  
Sequence: 1 ARVYTGLEFDTDFLKFVN.....VSLVDNDYINGDLEAFVADF 408

#### Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 424239 segs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+P2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09720451/runat\_05022003\_074810\_2410/app\_query.fasta\_1.583  
-DB=Published\_Applications\_NA -QPM=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09720451 -CGCN\_1.1.26 @runat\_05022003\_074810\_2410  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

#### Database :

Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCIT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCITUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	Score	Match	length	ID	Description
No.						
1	1777	81.3	1994	9	US-10-113-852A-3	Sequence 3, Appl1
2	979	44.8	2219	9	US-09-931-795-3	Sequence 3, Appl1
3	979	44.8	2219	10	US-09-728-910-3	Sequence 3, Appl1
4	979	44.8	2220	9	US-09-931-795-1	Sequence 1, Appl1

5	979	44.8	2220	10	US-09-728-910-1	Sequence 1, Appl	
C	950	43.4	832	10	US-09-770-445-690	Sequence 690, Appl	
7	778	35.6	3082	9	US-10-113-852A-1	Sequence 1, Appl	
8	775	35.4	574	10	US-09-734-017A-61	Sequence 61, Appl	
9	398	18.2	313	10	US-09-878-574-3765	Sequence 3765, Ap	
10	390	17.8	264	10	US-09-923-876-5596	Sequence 5596, Ap	
11	389	17.8	265	10	US-09-923-876-476	Sequence 476, App	
12	284	13.0	181	10	US-09-923-876-5753	Sequence 5753, Ap	
13	274	12.5	216	10	US-09-923-876-1503	Sequence 1503, Ap	
14	255	11.7	159	10	US-09-294-0938-4653	Sequence 4653, Ap	
15	213	9.7	409	10	US-09-734-017A-63	Sequence 63, Appl	
16	168	7.7	640681	9	US-09-780-888-1	Sequence 1, Appl	
17	158	7.2	378	9	US-09-738-626-2379	Sequence 2379, Ap	
18	157	7.2	1351	10	US-09-919-935-1	Sequence 1, Appl	
19	138	6.3	4848	9	US-09-943-702-1	Sequence 1, Appl	
20	138	6.3	4848	9	US-09-943-702-4	Sequence 4, Appl	
21	138	6.3	4848	9	US-09-943-702-6	Sequence 6, Appl	
22	100	4.6	38584	9	US-10-114-170-50	Sequence 50, Appl	
23	97.5	4.5	8220	9	US-10-153-273-11	Sequence 11, Appl	
C	24	93.5	1900	10	US-09-911-888-22	Sequence 22, Appl	
25	93.5	4.3	1900	10	US-09-911-888-24	Sequence 24, Appl	
26	91	4.2	2080	9	US-10-003-382-1	Sequence 1, Appl	
C	27	89.5	4.1	7053	10	US-09-070-927A-78	Sequence 78, Appl
28	88	4.0	1182	9	US-10-071-338-6	Sequence 6, Appl	
29	88	4.0	3364	10	US-09-925-300-665	Sequence 665, Appl	
C	30	88	4.0	7193	9	US-10-071-338-1	Sequence 1, Appl
C	31	87.5	4.0	1610	10	US-09-954-456-92	Sequence 92, Appl
C	32	87.5	4.0	1610	10	US-09-967-768A-219	Sequence 219, Appl
33	87.5	4.0	2393	10	US-09-964-824A-254	Sequence 254, Appl	
34	87.5	4.0	2433	9	US-09-738-626-2317	Sequence 2317, Ap	
35	87.5	4.0	4765	12	US-10-090-624-5	Sequence 51, Appl	
36	87	4.0	2956	10	US-09-895-382-31	Sequence 31, Appl	
37	85.5	3.9	1329	9	US-09-738-626-2703	Sequence 2703, Ap	
C	38	85.5	3.9	1674	9	US-09-738-626-2702	Sequence 2702, Ap
C	39	85.5	3.9	6449	9	US-09-938-842A-69	Sequence 69, Appl
C	40	85	3.9	963	10	US-09-070-927A-547	Sequence 547, Appl
C	41	85	3.9	3826	9	US-09-927-091-3	Sequence 3, Appl
C	42	85	3.9	6544	9	US-10-004-381-40	Sequence 40, Appl
C	43	85	3.9	6688	9	US-10-004-381-27	Sequence 37, Appl
44	84.5	3.9	1846	10	US-09-823-356-24	Sequence 24, Appl	
45	84.5	3.9	2070	12	US-10-044-090-377	Sequence 377, Appl	

#### ALIGNMENTS

RESULT 1  
US-10-113-852A-3

Sequence 3, Application US/10113852A  
Publication No. US20020192784A1

#### GENERAL INFORMATION:

APPLICANT: Applling, Dean R.  
APPLICANT: Hanson, Andrew D.  
APPLICANT: Raymond, Rhonda R.  
APPLICANT: Roje, Sanja  
TITLE OF INVENTION: Biosynthesis of S-adenosyl Methionine in a Recombinant Yeast S.  
FILE REFERENCE: 119927-1033  
CURRENT APPLICATION NUMBER: US/10/113, 852A  
PRIOR FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: 60/280,333  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3  
LENGTH: 1994  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (63)..(1841)  
OTHER INFORMATION: DNA construct  
US-10-113-852A-3

Alignment Scores:

Pred. No.: 2,16e-197 Length: 1994  
 Score: 1777.00 Matches: 319  
 Percent Similarity: 89.41% Conservative: 44  
 Best Local Similarity: 78.57% Mismatches: 43  
 Query Match: 81.25% Indels: 0  
 DB: 9 Gaps: 0

US-09-720-451-6 (1-408) x US-10-113-852A-3 (1-1994)

Qy 1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20  
 Db 621 GAGATTTGATGTGACCTGCTTCTATGATGCTGATGATGATGATGATGATGATGAT 680  
 Qy 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn 40  
 Db 681 GATGTGCGCAAAATGCGGATTAATGTCATGTCCTGGAATTAAGCTTATTCCTAAC 740  
 Qy 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProIleAspIleMet 60  
 Db 741 TACAAAGGGGTTCTTGCGATGCTGCTGTTCTGTAAAGACCAAGATACCCGCTGACT 800  
 Qy 61 AlaAlaLeuGlnProIleLysAspAsnGlnIleValIleValIleValIleValIle 80  
 Db 801 GCTGCTTACAGGCTTATTAAGATATACAGAGGCTGTAAAGCTTATGGAATTCCTT 860  
 Qy 81 GlyThrGlnMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThr 100  
 Db 861 GCAACGAAATGTGCAAAAGATTTGGCCCAAGGATGATCTCCCTGCTCTACACA 920  
 Qy 101 LeuAsnMetGlnLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlnIle 120  
 Db 921 TTGAACTGCAAAATCACCTATGCGATATTAAAGAACTGTGCTGATGATGATGAT 980  
 Qy 121 LysValSerArgSerLeuProTyrArgTyrProIleAsnValPheArgValLysGlnAsp 140  
 Db 981 AAAATTTCTGCTCTCTACCTTGAGAGCCCTGCAAAATGTTTCCGCTACTAAGAGAT 1040  
 Qy 141 ValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGlyTyr 160  
 Db 1041 GTTGCGCCCAATTTCTGGGCAAAAGCCCAAGACCTACATATCTAACAAAGGGCTCG 1100  
 Qy 161 AspGlnTyrProHisGlyArgTyrPalaAspSerCysAsnProSerTyrGlyAlaLeuSer 180  
 Db 1101 AATGACTCCACATGAGAGCTTGCGGTGATTCACACAGTGCAGATACATACCTTGC 1160  
 Qy 181 AspTyrGlnPheMetArgProArgPalaArgAspLysLeuValGlnIleTyrPalaVal 200  
 Db 1161 GATTATAGTTTCCGCCCAAGAGAGCTGACAGAGAGCTTCAGCAAGATGCGTCATC 1220  
 Qy 201 ProLeuLysSerValGlnAspIleTyrGlnArgPheArgLeuTyrCysLeuGlyLysLeu 220  
 Db 1221 CCACTGAAAGCAATTAAGATTTTCAGAGAAATTCAAAGAGCTCTCATTTGAAACTTA 1280  
 Qy 221 ArgSerAsnProTyrSerGlnLeuAspGlyLeuGlnProGlnThrLysIleIleAsnGln 240  
 Db 1281 AAAAGCAGCCCATGCTGATTAATGATGAGCTCCAGCAGAGCAAGATCATTAATGAG 1340  
 Qy 241 GlnLeuGlnLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260  
 Db 1341 CAAGCTCGAAGAAATCACTCAAGCGTTTCCGACATTAATACCAACCATAGTCATAT 1400  
 Qy 261 GlyLysLysSerAspSerProThrValGlyTyrPalaGlyProGlyLysIleTyrGln 280  
 Db 1401 GCAAGCCAAATCCGATTTCTCAGCTATGATGATGGGTGCTCTGCTGCTGCTGCT 1460  
 Qy 281 LysAlaTyrValGlnPhePheCysSerLysGlnLysLeuAspAlaLeuValAspLysCys 300  
 Db 1461 AAACCTTATCTGAGATGTTCTGCTCAAGGATAGTACACACTTGTGGGAAATGCTC 1520  
 Qy 301 LysAspArgThrSerLeuThrTyrIleAlaValAsnLysAspGlySerTyrLysSerAsn 320  
 Db 1521 AAACCTTTCTCTATCACTACATGCGCGTGAACCAATCAGAGATTTGGTATCAAC 1580

Qy 321 ValGlyGlnThrAspValAsnAlaValThrTyrPalaValPheProAlaLysGlnIleIle 340  
 Db 1581 ACCGCTGAGTCCGATGCTAAATGCAAGTACTTGGGAGGTCTTCCAGCTAAGAGGTATT 1640  
 Qy 341 GlnProThrIleValAspProValSerPheAsnValTyrLysAspGlnAlaPheGlnIle 360  
 Db 1641 CAACGACGATCGTGGATCCAGCGTTTCAAGTGTGAAAGCAAGCGTTTGAGATT 1700  
 Qy 361 TyrSerArgGlyTyrPalaSerLeuTyrProGlnLysAspGlnAlaSerArgLysLeuValGln 380  
 Db 1701 TGGTCAGAGAGCTGGCTTAATCTGTACCCAGAGATGACCTTCTAAGAAAGTGTGAG 1760  
 Qy 381 GlnValGlyLysSerIlePheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400  
 Db 1761 GAGGTGAGACAGCTACTATTGTTGTAAGCTTAGTGACAAACAAATTAATCAATGCTGAT 1820  
 Qy 401 LeuPheAlaValPheAla 406  
 Db 1821 AATCTCTGCTGCTTGGCT 1838

#### RESULT 2

US-09-931-795-3

Sequence 3, Application US/09931795

Publication No. US2002019821A1

GENERAL INFORMATION:

APPLICANT: ROZEN, Rima

TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETHETRAHYDROFOLATE

TITLE OF INVENTION: REDUCTASE AND USES THEREOF

FILE REFERENCE: 04844/005003

CURRENT APPLICATION NUMBER: US/09/931,795

CURRENT FILING DATE: 2001-08-16

PRIOR APPLICATION NUMBER: US 09/592,595

PRIOR FILING DATE: 2000-06-12

PRIOR APPLICATION NUMBER: US 09/258,928

PRIOR FILING DATE: 1999-03-01

PRIOR APPLICATION NUMBER: US 08/738,000

PRIOR FILING DATE: 1997-02-12

PRIOR APPLICATION NUMBER: PCT/CA95/00314

PRIOR FILING DATE: 1995-05-25

PRIOR APPLICATION NUMBER: GB 9410620.0

PRIOR FILING DATE: 1994-05-26

NUMBER OF SEQ ID NOS: 18

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 3

LENGTH: 2219

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (13)...(1983)

US-09-931-795-3

#### Alignment Scores:

Pred. No.: 3.09e-104 Length: 2219  
 Score: 979.00 Matches: 188  
 Percent Similarity: 84.72% Conservative: 79  
 Best Local Similarity: 44.87% Mismatches: 138  
 Query Match: 44.76% Indels: 14  
 DB: 9 Gaps: 7

US-09-720-451-6 (1-408) x US-09-931-795-3 (1-2219)

Qy 1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20  
 Db 676 GCCGATTCATCATCATCCAGCACTTCTTGAAGCTGACACATCTCTCCGCTTGTGAAG 735  
 Qy 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn 40  
 Db 736 GCATGACCGACATGGGATCATCTGCCCCCATGCTCCCGGATCTTCCATCCAGGCG 795  
 Qy 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60  
 Db 796 TACCACTCCCTTCCGACAGCTTGTGAAGCTGTCCAGAGTGGAGTGCACAGAGATCAAG 855

```
QY 61 AAlaAlaLeuGluProIleuLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 80
Db 856 GAGCGATTGAGCAATCAAGAACAGATGGCTGCATCGCATAGGATCGAGCTG 915
QY 81 GAlTThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyr 99
Db 916 GCGGTGAGCGCTGTGCGAGAGCTTTCGGCAGAGCTTGAGTGGCCGAGCCCTCACTTCTAC 975
QY 100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlu 119
Db 976 ACCCTCAACCGCGAGATGGCTACCAACAGAGCTGTGAAGCCGCTTGAGATGTGACAG 1035
QY 120 SerLysValSerArgSerLeuProTTPArgArgProAlaAsnValPheArgValLysGlu 139
Db 1036 GACCCC---AGCGCTCCCTACCTCGGCTGTCAAGGCCACCCACCGCGGAGAGAA 1092
QY 140 AspValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159
Db 1093 GATGTACGCTCCATCTTCTGGCGCTCCAGCAAGAGTTACATCTACCGTACCCAGAG 1152
QY 160 TPAspGlnTyrProHisGlyArgTTPGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
Db 1153 TGGGAGAGTCCCTCAAGCGCGCTGAGCATCTCTCCCTGCGCTTGGGGAGACTG 1212
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197
Db 1213 AAGGACTACTACTCTTCTACCTGAGAGAGCAAGTCCCGCCACAGAGAGCTGTGAAGATG 1272
QY 198 TPAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1273 TGGGGGAGAGAGCTGACCAAGCAAGTGTCTGAAGTCTTGAAGTCTTGTCTTACCTCTCG 1332
QY 218 GAlTylsLeuAspSerAsp-----ProTyrSerGluLeuAspGly 230
Db 1333 GGGAAACCAACCGGAATGGTCACAAAGTGACTTGGCTCCCTCGGAAAGAT---GACCCC 1389
QY 231 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGlyLysIleAsnThrLysGlyPhe 250
Db 1390 CTGGGGCGCTGAGACGAGCTGTGAGAGAGAGAGCTGTGCGGGTGAACCGCGAGGCTATC 1449
QY 251 LeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrValGly 270
Db 1450 CTCACCATCACTCAACGCCCAACATCAACAGCGGAACCGCTCTCGACCCCATGTGTGGC 1509
QY 271 TPArgGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
Db 1510 TGGGGGCCCAAGGGGGCTATGTCTTCACAAAGCGCTACTTGAAGTTTTCATCTTCCCGC 1569
QY 291 GAlTylsLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
Db 1570 GAGACAGCGGAAGCACTTCTGCMAAGTGTGAGAGATACAGACTCGGGGTAAATATACAC 1629
QY 310 AlaValAsnLysAspGlySerTTPylsSerAsnValGlyGlnThrAspValAsnAlaVal 329
Db 1630 CTGTGCAATGTGAAGGGGAAACATCAACCAATGCCCTGAATCTGCACGCCAATGCTGTC 1689
QY 330 ThrTPArgValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
Db 1690 ACTTGGGAGATCTTCCCTGGGCGAGAGATCATCCAGCCACGATGTGATCCCGTCAGC 1749
QY 350 PheAsnValTPylsAspGluAlaPheGluIleTPylsSerArgLysTTPAlaSerLeuTyr 369
Db 1750 TTATGTTCTGAGAGAGAGAGCGCTTGGCTGTGATGGAGACGGGTGGGAAACGTGTAT 1809
QY 370 ProGluAspGluAlaSerArgLysLeuValGluGlnValGlyLysSerHisPheLeuVal 389
Db 1810 GAGGAGAGATCCCGCTCCGACCATCATCACTACATCAACGACAACTACTTCTGTGTC 1869
QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
Db 1870 AACCTGTGAGCAATGACTCTCCACATGTGACAACTGCTCTGGCAGAGTGTGGAAGAC 1926
```

```
RESULT 3
US-09-728-910-3
; Sequence 3, Application US/09728910
; Patent No. US20010025030A1
; GENERAL INFORMATION:
; APPLICANT: Rozen, Rima
; APPLICANT: Sekhon, Jaspreet
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/006001
; CURRENT APPLICATION NUMBER: US/09/728,910
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(1983)
US-09-728-910-3

Alignment Scores:
Pred. No.: 3,09e-104 Length: 2219
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Indels: 14
DB: Gaps: 7

US-09-720-451-6 (1-408) x US-09-728-910-3 (1-2219)
QY 1 AAlaAlaLeuGluProIleuLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 20
Db 676 GCGGATTCATCACTGACGAGCTTCTTGTGAGCGTGAACATCTTCCGCTTGTGAG 735
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValIleProGlyIleMetProIleAsnAsn 40
Db 736 GATGAGCGAGATGGGATGATCACTGCGCCATGCTGCCGGATGTTCCATCCAGAGCG 795
QY 41 TPylsGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60
Db 796 TACCACTCCCTTGGGAGACTGTGAGAGCTGTCCAACTGGAAGTGTCCACAGAGATCAAG 855
QY 61 AAlaAlaLeuGluProIleuLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 80
Db 856 GAGCGATTGAGCAATCAAGAACAGATGGCTGCATCGCATAGGATCGAGCTG 915
QY 81 GAlTThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyr 99
Db 916 GCGGTGAGCGCTGTGCGAGAGCTTTCGGCAGAGCTTGAGTGGCCGAGCCCTCACTTCTAC 975
QY 100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlu 119
Db 976 ACCCTCAACCGCGAGATGGCTACCAACAGAGCTGTGAAGCCGCTTGAGATGTGACAG 1035
QY 120 SerLysValSerArgSerLeuProTTPArgArgProAlaAsnValPheArgValLysGlu 139
Db 1036 GACCCC---AGCGCTCCCTACCTCGGCTGTCAAGGCCACCCACCGCGGAGAGAA 1092
QY 140 AspValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159
Db 1093 GATGTACGCTCCATCTTCTGGCGCTCCAGCAAGAGTTACATCTACCGTACCCAGAG 1152
QY 160 TPAspGlnTyrProHisGlyArgTTPGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
Db 1153 TGGGAGAGTCCCTCAAGCGCGCTGAGCATCTCTCCCTGCGCTTGGGGAGACTG 1212
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197
```

```
Db 1213 AAGACTACTACCTTCTTACTGAAAGCAAGTCCCAAGAGAGAGACTGCTGAAGATG 1272
Qy 198 TTPAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1273 TGGGGGAGAGAGCTGACACAGTGAAGCAAGTCTTGAAGCTTTGTTTACTTACCTTCG 1332
Qy 218 GLYLYSLeuArgSerAsn-----ProTyrSerGluLeuAspGly 230
Db 1333 GGAGAACCAACCGGAATGTCACAAAGTACTGCTGCTGCTGCAACAT---GAGCCC 1389
Qy 231 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe 250
Db 1390 CTGGGGCTGAGACACAGCTCTGAGAGAGAGCTGCGGGGGAAGCCGACGAGGCGATC 1449
Qy 231 LeuThrIleAsnSerGlnProAlaValAsnGluLysSerAspSerProThrValGly 270
Db 1450 CTCACATCAACATCACAGCCCAACATCAACGGGAAGCCGCTCCGACCCATCGTGGGC 1509
Qy 271 TTPGlyLysProGlyLysTyrValTyrGluLysAlaTyrValGluPhePheCysSerLys 290
Db 1510 TGGGGCCCAAGCGGGGCTATGTCCTCCAGAAAGCGCTACTAGAGTTTTCACCTCCGC 1569
Qy 291 GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
Db 1570 GAGACAGCGAGACACTTCTGCAAGTCTGAGAAAGTACAGACTCCGGGTTAATTAACAC 1629
Qy 310 AlaValAsnLysAspGlySerTyrLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
Db 1630 CTTCATCATGTGAAGGTGAAACATCAACATGCCCCCTGACATGCAAGCCGATGCTGTC 1689
Qy 330 TTTTTPGlyLysPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
Db 1690 ACTGGGGGCACTCTCCCTGGGGGAGAGATCAATCCAGCCACCGAAGGATCCCGTCAAGC 1749
Qy 350 PheAsnValIlePheProAlaLysAspGluIlePheGluIleTyrSerArgLysTyrAlaSerLys 369
Db 1750 TTTCATGTTCTGGAAGAGACAGGCTTTGCGCTGTGATGAGCGGAGGGAACGCTGAT 1809
Qy 370 ProGluAspGluAlaSerArgLysLeuValGluGluValGlyLysSerHisPheLeuVal 389
Db 1810 GAGGAGGAGTCCCGCTCCGACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1869
Qy 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
Db 1870 AACCTGTGGACATGACTTCCCACTGAGCAACTGCTGTGGCAGGTGGTGAAGAC 1926

RESULT 4
; Sequence 1, Application US/09931795
; Publication No. us20020198211A1
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/005003
; REDUCTASE AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/931,795
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 09/592,595
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/738,000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: PCT/CA95/00314
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1980)
; US-09-931-795-1

Alignment Scores:
Pred. No.: 3 09e-104 Length: 2220
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44./68 Indels: 14
Gaps: 7

US-09-720-451-6 (1-408) x US-09-931-795-1 (1-2220)
Qy 1 AlaArgValIleValThrGlnLeuPheTyrAspThrIlePheLeuLysPheValAsn 20
Db 676 GCCGATTTCATCATCAACGACGAGCTTTCTTGAGGCTGACACATCTTCCGCTTGTGAAG 735
Qy 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
Db 736 GCATGACACGACATGGCGATTCATCCCTGCCCATCCCTCCCGGATCTTCCATCCAGGCG 795
Qy 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60
Db 796 TACCACTCCCTTGGCAGCTTGTGAAGCTGTCCAAAGCTGGAGGTGCCACAGAGATCAAG 855
Qy 61 AlaAlaLeuGluProIleLysAspAsnGluGluAlaValAlaLysAlaTyrGlyIleHisLeu 80
Db 856 GACGTGATGAGCCAAATCAAAAGACATGCTCCATCCGCAACATATGATCGACGCTG 915
Qy 81 GlyThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrIleHisLeuTyr 99
Db 916 GCCCTGAGCCTGTGCGAGAGCTTGTGCCAGTGGCTTGGTGGCAGGCTCCATCTTAC 975
Qy 100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetLsnLeuGluIleGlu 119
Db 976 ACCCTCAACCGCAGATGCTACACAGAGGTGTGAAGCGCTGGAGTGGAGCTGAG 1035
Qy 120 SerLysValSerArgSerLeuProTyrArgArgProAlaAsnValPheArgValLysGlu 139
Db 1036 GACCCC---AGGCGTCCCTTACCTGCGGCTCTCACTGCTCCACCCCAAGCGCGAAGGAA 1092
Qy 140 AspValArgProIlePheThrPheAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159
Db 1093 GATGTAGTCCCATCTTCTTGGGCTCCGACCAAAAGTTATCATCTTACCGTACCAGAGAG 1152
Qy 160 TTPAspGluTyrProHisGlyArgTTPGlyAspSerCysAsnProSerGlyValAlaLeu 179
Db 1153 TGGAGCAGTTCCTTACAGCGCGCTGGGGCAATTCCTCTTCCCTCCTTGGGAGCTG 1212
Qy 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu 197
Db 1213 AAGACTACTACCTTCTTACTGAAAGCAAGTCCCAAGAGAGAGTGTGAAAGT 1272
Qy 198 TTPAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1273 TGGGGGAGAGAGCTGACACAGTGAAGCAAGTCTTGAAGCTTTGTTTACTTACCTTCG 1332
Qy 218 GLYLYSLeuArgSerAsn-----ProTyrSerGluLeuAspGly 230
Db 1333 GGAGAACCAACCGGAATGTCACAAAGTACTGCTGCTGCTGCAACAT---GAGCCC 1389
Qy 231 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe 250
Db 1390 CTGGGGCTGAGACACAGCTCTGAGAGAGAGCTGCGGGGGAAGCCGACGAGGCGATC 1449
Qy 251 LeuThrIleAsnSerGlnProAlaValAsnGluLysSerAspSerProThrValGly 270
Db 1450 CTCACATCAACATCACAGCCCAACATCAACGGGAAGCCGCTCCGACCCATCGTGGGC 1509
Qy 271 TTPGlyLysProGlyLysTyrValTyrGluLysAlaTyrValGluPhePheCysSerLys 290
```

Db 1510 TGGGGCCCCAGGGGGCTATGTCTCCAGAAAGCCCTACTAGAGTTTTCACCTCCCGC 1569  
QY 291 GILULYSLEUASPALALEUVALASPLYSGLYSASP--ArgThrSerLeuThrTyrMet 309  
Db 1570 GAGACAGCGGAGCACTTCTCAAGTGTCTGAAGAGTACAGAGCTCCGGGTAAATTTCACAC 1629  
QY 310 AlAlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329  
Db 1630 CTGTGCAATGTGAAGGTGAAACATCATCCAAATGCCCTGACAGCGCCGATGCTGTGC 1689  
QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349  
Db 1690 ACTTGGGGCATCTCTCCGTGGGAGAGATCATCCAGCCCAACCGTAGTGATGCCGTCAGC 1749  
QY 350 PheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgLysTrpAlaSerLeuTyr 369  
Db 1750 TTCATGTCTCGAAGAGACAGAGCCCTTTCCTGTGATGAGCGGTGGGAAAGCTGTAT 1809  
QY 370 ProGluAspGluAlaSerArgLysLeuValGluGluValGlySerHisPheLeuVal 389  
Db 1810 GAGAGAGAGTCCCGCTCCCGCACATCATCCAGTACATCCAGCAACTACTCTGTGTC 1869  
QY 390 SerLeuValAspAsnAspTyr--IleAsnGlyAspLeuPheAlaValPheAlaAsp 407  
Db 1870 AACCTGTGTGACATGACTTCCCACTGGACACAACTGCTTGTGGCAGGTGTGGAAGAC 1926

RESULT 5  
US-09-728-910-1  
; Sequence 1, Application US/09728910  
; Patent No. US20010025030A1  
; GENERAL INFORMATION:  
; APPLICANT: Rozen, Rima  
; APPLICANT: Sekhon, Riaspreet  
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETERAHYDROFOLATE  
; FILE REFERENCE: 04844/006001  
; CURRENT APPLICATION NUMBER: US/09/728,910  
; CURRENT FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: US 09/258,928  
; PRIOR FILING DATE: 1999-03-01  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2220  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1980)  
US-09-728-910-1

Alignment Scores:  
Pred. No.: 3,09e-104 Length: 2220  
Score: 979.00 Matches: 188  
Percent Similarity: 63.72% Conservative: 79  
Best Local Similarity: 44.87% Mismatches: 138  
Query Match: 44.76% Indels: 14  
Gaps: 7

US-09-720-451-6 (1-408) x US-09-728-910-1 (1-2220)

QY 1 AlArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20  
Db 676 GCCGATTTCATCATCATCGAAGCTTTCTTTGAGCTGACACATCTTCGCGTTTGTAAG 795  
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40  
Db 736 GCATGCACCGACATGGGCTACCTTGCCCACTGCCGCGGATCTTCCATCCACAGGC 795  
QY 41 TyrIleGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60  
Db 796 TACACATCCCTTGGCAGAGCTGTGAAGCTGTCCAAAGCTGGAGGTGCCACAGAGATCAAG 855

QY 61 AlAlaLeuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 80  
Db 856 GAGCTGATTAGCCATCAAGACACAGCATGTGCGCATCCGCACTGGCATGTGAGACTG 915  
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGly--IleLysThrLeuHisLeuTyr 99  
Db 916 GCCGTGACCTGTGCCAGAGAGCTTGTGGCCAGTGGCTGTGGAGCGCTTCACCTTCAC 975  
QY 100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlu 119  
Db 976 ACCCTCAACCGCAGATGGCTACACAGAGTGTGAAGCCCTGGGATGTGACAGAG 1035  
QY 120 SerLysValSerArgSerLeuProTrpArgArgProLysAsnValPheArgValLysGlu 139  
Db 1036 GACCCC--AGGGGTCCCTACCCCTGGGCTCTCACTGGCCCAACCCCAAGCCGAGAGAA 1092  
QY 140 AspValArgProIlePheThrPheAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159  
Db 1093 GATGTACGTCCCATCTTGTGGGCTCCAGACCAAGAGTACATCTTACCGTACCCAGAG 1152  
QY 160 TrpAspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 179  
Db 1153 TGGAGCAGATTCCCTGAACGGCCGTGGGCAATTCCTCTCCCTGCTGGGGAGACTG 1212  
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLeuValGluGlu 197  
Db 1213 AAGGACTACTACTCTTCTTACCTGGAAGAGCAAGTCCCAAGAGAGAGACTGTCAAGATG 1272  
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217  
Db 1273 TGGGGGAGAGAGCTGACAGATGAACAGCAAGTGTGGAAGTCTTGTCTTACCTCTCG 1332  
QY 218 GlyLysLeuArgSerAsn-----ProTrpSerGluLeuAspGly 230  
Db 1333 GGAACACCAACCGGATGGTCAAAAGTCACTTCCCTCCCTGGAGAGAT--GAGCC 1389  
QY 221 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe 250  
Db 1390 CTGGCGCTGAGACAGCCGTGCTGAAGAGAGAGTGTGGGGTGAACCGCAGGAGCATC 1449  
QY 251 LeuThrIleAsnSerGlnProAlaValAlaAsnGlyLysSerAspSerProThrValGly 270  
Db 1450 CTCACATCAACTACACAGCCCAACATCAAGGGAAGCGTCTCCGACCCCATGCTGGGC 1509  
QY 271 TrpGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290  
Db 1510 TGGGGCCCCAGGGGGCTATGTCTCCAGAAAGCCCTACTTGAAGTTTTCACCTCCGC 1569  
QY 291 GILULYSLEUASPALALEUVALASPLYSGLYSASP--ArgThrSerLeuThrTyrMet 309  
Db 1570 GAGACAGCGGAGCACTTCTCAAGTGTCTGAAGAGTACAGAGCTCCGGGTAAATTTCACAC 1629  
QY 310 AlAlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329  
Db 1630 CTGTGCAATGTGAAGGTGAAACATCATCCAAATGCCCTGACAGCGCCGATGCTGTGC 1689  
QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349  
Db 1690 ACTTGGGGCATCTCTCCGTGGGAGAGATCATCCAGCCCAACCGTAGTGATGCCGTCAGC 1749  
QY 350 PheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgLysTrpAlaSerLeuTyr 369  
Db 1750 TTCATGTCTCGAAGAGACAGAGCCCTTTCCTGTGATGAGCGGTGGGAAAGCTGTAT 1809  
QY 370 ProGluAspGluAlaSerArgLysLeuValGluGluValGlySerHisPheLeuVal 389  
Db 1810 GAGAGAGAGTCCCGCTCCCGCACATCATCCAGTACATCCAGCAACTACTCTGTGTC 1869  
QY 390 SerLeuValAspAsnAspTyr--IleAsnGlyAspLeuPheAlaValPheAlaAsp 407  
Db 1870 AACCTGTGTGACATGACTTCCCACTGGACACAACTGCTTGTGGCAGGTGTGGAAGAC 1926

RESULT 6

```

Db      470  ATTACTTACATGGCTGTGACCAAGAGACAGTGGGATTCGAAACACTGCCCAAGCTGAT 411
OY      326  ValAsnAlaValThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleVal 345
Db      410  GTGATATGCTTAACTTGGGGAGATTTCCTCCGCTAAGGAATCATTCACCAACCATTTGTC 351
OY      346  AspProValSerPheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgGlyTrp 365
Db      350  GATCCCGCTGACTTCAACGCTGGAAGCATGACGATTTGGACTTGGTCAAGAAGCTGG 291
OY      366  AlaSerIleuTrpGluLysAspGluAlaSerArgLysIleuValGluGluValGlyLysSer 385
Db      290  GCTACTCTGTACCCAGAACCCACCTTCCAGAACTTGCTCGAGGAGTGAAGAACAGC 231
OY      386  HisPheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAspLeuPheAlaValPhe 405
Db      230  TACTATTGGTAGAGCTTGTGGAGAACGATATACATCAATGCGTACATATTCGGGCTCTT 171
OY      406  AlaAsp 407
Db      170  GCTGAT 165

RESULT 7
US-10-113-852A-1
; Sequence 1, Application US/10113852A
; Publication No. US20020192784A1
; GENERAL INFORMATION:
; APPLICANT: Appling, Dean R.
; APPLICANT: Hanson, Andrew D.
; APPLICANT: Raymond, Rhonda R.
; APPLICANT: Kojer, Sanja
; TITLE OF INVENTION: Biosynthesis of S-Adenosyl Methionine In a Recombinant Yeast S
; FILE REFERENCE: 119927-1033
; CURRENT APPLICATION NUMBER: US/10/113, 852A
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 60/280,333
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3082
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (826)..(2625)
; OTHER INFORMATION: DNA construct
US-10-113-852A-1

Alignment Scores:
Pred. NO.: 1.43e-80 Length: 3082
Score: 778.00 Matches: 168
Percent Similarity: 59.37% Conservative: 76
Best Local Similarity: 40.88% Mismatches: 147
Query Match: 35.57% Indels: 20
                Gaps: 11

US-09-720-451-6 (1-408) x US-10-113-852A-1 (1-3082)
OY      4  IleValThrGluLeuPheTyrAspThrAspIlePheLeuLysPheValAsnAspCysArg 23
      |||:|||||:|||||:|||||  ||  |||  ::  ::  |||
Db      1378  ATCATCATCATGATGTTTATTCATGTGTGATTAATTACTCAACACTGGTGTCCCAAGTTAA 1437
OY      24  GlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsnTyrLysGly 43
      |||:|||||:|||||:|||||:|||||  |||
Db      1438  GCTCGGGGCGATGAGCGTCCCATTAATTCGCCGAGATGCGGATCACTACTACGCGGCGC 1497
OY      44  PheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMetAlaIleVal 63
      |||:|||||:|||||:|||||  |||
Db      1498  TTCTTGGAGAAAGATCCAAATGGGGCCAAATTCATCCTCTCAACATTTCTGTGTCGCGATTG 1557
OY      64  GlnProIleuLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeuGlyThrGlu 83

```

Db	1558	GATCGTATCAAGAGCATTGACGATGGTGGTCCGTGATATCGAACAATTGATGTCGTGAA	1617
Qy	84	MetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLysLeuThrLeuAsn	102
Db	1618	ATGTGTCAAAAATGTCGCAAGATGGTATACGTTTCCACTTCACATTCACATTCACCATGAC	1677
Qy	103	MetClnLysSerIleLeuAlaIleLeuMetAsnLeuGlyLeuIle---GlnGlnSerLys	121
Db	1678	TTGGAAATAAACGCCCTCTCATGATTTTGGAAAGATGATGAACATTCTCACTAGCAATCAAG	1737
Qy	122	ValSerArgSer-----LeuProTyrArgArgProAlaAsnValPheArgVal	137
Db	1738	TTCAATGCACATCCATTGGCGCGTGGCCATGGAGAAATCTTTAAATCCAAAGCGTAAA	1797
Qy	138	LysGlnAspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThr	157
Db	1798	AACGAGAACTCAGACCTATCTCTTGGAAAGAGAACCTTACTCTCATATGTCGCAAGAAC	1857
Qy	158	IleGlyIrr-----AspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSer	175
Db	1858	TCTCAATAGGGCGCGTGAGCAAGATCCCCCAACGGATAGATTGGTGATTCCTCTCTCCG	1917
Qy	176	TyrGlyAlaLeu-----SerAspTyrGlnPheMetArgProArgAlaArgAsp	191
Db	1918	TTCCGCTACTTACGTATGCTGTGGTGCAGAC-----TTGATCAGCAATCAAGCG-----	1965
Qy	192	LysLysLeuValGlnGlnTrrPalaValProLeuLysSerValGlnAspIleArgGlnArg	211
Db	1966	AACAAATGCTCGAATATATGTCACCCCT---ACTTCATCAACAGCAGCTGCCTTCTTG	2022
Qy	212	PheArgLeuTyrCysLeuGlyLysLeuArgSerAsnProTrrSerGlnLeuAspLysLeu	231
Db	2023	GTCATCAACTACTTGATGGATGGAAACTTGAGAGTGTTCACCTGGAGATGATATCCCC---ATC	2079
Qy	232	GlnProGlnThrLysIleIleAsnGlnGlnLeuGlnLysIleAsnThrLysGlyPheLeu	251
Db	2080	AATGATATAATTAATCCAAACAAAGACACACTGGATTGACCTTAACACAGCATTCATATCATC	2139
Qy	252	ThrIleAsnSerGlnProAlaValAsnGlyLysLysSerAspSerProThrValGlyTyr	271
Db	2140	ACTATAACTCTCAACTCAAGTCAACGCAAGCGCATTAGTCCAAATGACAAATTCATCGTTGG	2199
Qy	272	GlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGlnPhePheCysSerLysGln	291
Db	2200	GGACCCAGAGATGGTTACGTTTACCAAGACATATTTGGAATTTATGTGCCCAAGACT	2259
Qy	292	LysLeuAspAlaLeuValAspLysCysLysAspArgThrSerLeuThrTyrMetAlaVal	311
Db	2260	AAGTGGCCCAAGTTGATGTGACACTTGAAAAAACAAGAATCTTGACTACTCTGCCATC	2319
Qy	312	AsnLysAspGlySerTrrPlysSerAsn---ValGlyGlnThrAspValAsnAlaValThr	330
Db	2320	GACTCTCAAGCTGACCTGCTATATCATATCCAGACATCCAACTCAACACGCTGTACT	2379
Qy	331	TrrGlyValPheProAlaLysGlnIleIleGlnProThrIleValAspProValSerPhe	350
Db	2380	TGGGGTATTTTCCCGGCAAGAAATTTCTCAACCTTCACTTGTGAGAAATTTTCGTC	2439
Qy	351	AsnValTrrPlysAspGlnAlaPheGlnIleTrrPserArgGlyTrrPalaSerLeuTyrPro	370
Db	2440	TTACGCTGGAGAGGAGGAGTCTTATCATATCTTGAAATGAATGAACTAAACATGAATAAA	2499
Qy	371	GlnAspGlnAla---SerArgLysLeuValGlnGlnValGlyLysSerHisPheLeuVal	389
Db	2500	TACATATAACCGCATTAATGCCCAATTCATTCACCTCCATCTTGATGACGATTAACGCTTGGTC	2559
Qy	390	SerLeuValAspAsnAspTyrIleAsnGlyAsp	400
Db	2560	AATATGTGTGACATATGCTACATTTCTCCAGAT	2592

RESULT 8  
US-09-734-017A-61

```

: Sequence 61 Application US/09734017A
: Patent No. US20020142422A1
: GENERAL INFORMATION:
: APPLICANT: Lerchl, Jens
: APPLICANT: Renz, Andreas
: APPLICANT: Ehrhardt, Thomas
: APPLICANT: Reindl, Andreas
: APPLICANT: Cirpus, Petra
: APPLICANT: Bischoff, Friedrich
: APPLICANT: Frank, Markus
: APPLICANT: Freund, Annette
: APPLICANT: Duwendig, Elke
: APPLICANT: Schmidt, Ralf-Michael
: TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved in
: TITLE OF INVENTION: the
: TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides and
: TITLE OF INVENTION: nucleosides
: FILE REFERENCE: BASE-NAE-1331-99-US
: CURRENT APPLICATION NUMBER: US/09/734,017A
: CURRENT FILING DATE: 2000-12-12
: PRIOR APPLICATION NUMBER: 60/171,100
: PRIOR FILING DATE: 1999-12-16
: NUMBER OF SEQ ID NOS: 87
: SOFTWARE: Patentin Ver. 2.1/WordPerfect
: SEQ ID NO 61
: LENGTH: 574
: TYPE: DNA
: ORGANISM: Physcomitrella patens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(573)
: OTHER INFORMATION: 86_ppprot1_094_g10
US-09-734-017A-61

```

[illegible]

```
Db 427 CCAATTTCGGGCAACGACCTAGAGTACATTTCAGAACCAACGAGTGGACGAT 486
Qy 163 TyrProHisIlyArGTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyr 182
Db 487 TTCCCTCGTGAAGGTGGGAGATAGGCC-ACCCCTGCTTACGCGAGCTTCAGCGATCAT 545
Qy 183 GluPheMetArpProArg 188
Db 546 CAGTTCCACGACGAGAAG 563

RESULT 9
US-09-878-574-3765
; Sequence 3765, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3765
; LENGTH: 313
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-006-Q1-B1-C9
US-09-878-574-3765

Alignment Scores:
Pred. No.: 8,89e-38 Length: 313
Score: 398.00 Matches: 78
Percent Similarity: 94.12% Conservative: 2
Best Local Similarity: 91.76% Mismatches: 5
Query Match: 18,208 Indels: 0
Gaps: 0

US-09-720-451-6 (1-408) x US-09-878-574-3765 (1-313)
Qy 322 GtGtGtThAspValAsnAlaValThrTrpGlyValPheProAlaLysGluIleIleGln 341
Db 1 GGTCAACGACGACTTAACTGCTGACATGGGGGTCTTCCAGCTAAGAGATAATTCAA 60
Qy 342 ProThrIleValAspProValSerPheAsnValTrpLysAspGluAlaPheGluIleTrp 361
Db 61 CCAACCATGTGTAGATCCTGTGACGCTTCATGTATGAGAGATGACCATTTGAATTTGG 120
Qy 362 SerArgGlyTrpAlaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGluGlu 381
Db 121 TCAAGATGATCGGCAAGCTGTACCTCGAGGATTAAGCATCCAGAAATGTGTGAAGAG 180
Qy 382 ValGlyGlySerHisPheLeuValSerLeuValAspAsnAspTyrTrpIleAsnGlyLeu 401
Db 181 GTTGGGGGACGACCTTCTTGGTACTTGTGTCGACAAATTATTAATCATGATGATCCT 240
Qy 402 PheAlaValAlaPheAla 406
Db 241 TTGGCGCTCTTGCA 255

RESULT 10
US-09-923-876-5596
; Sequence 5596, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laljundi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
```

```
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5596
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457310H1
; NAME/KEY: unsure
; LOCATION: 81, 183
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5596

Alignment Scores:
Pred. No.: 5,86e-37 Length: 264
Score: 390.00 Matches: 72
Percent Similarity: 87.50% Conservative: 5
Best Local Similarity: 81.82% Mismatches: 10
Query Match: 17,83% Indels: 1
Gaps: 0

US-09-720-451-6 (1-408) x US-09-923-876-5596 (1-264)
Qy 116 LeuIleGluGluSerLysValSerArgSerLeuProTrpAlaArgProAlaAsnValPhe 135
Db 3 TTAATTGAGAGATCCAAAGGTTTCAAGGCATTAACCTTGAGGCCAGGACGACTAAAGTTTC 62
Qy 136 ArgValLysGluAspVal-ArgProIlePheTrpAlaAsnArgProLysSerTyrIleSe 155
Db 63 CCGTTAAAGAGAGATGTTNCACACTATATTTGTGGCCCAAGACCAAGACTATCTTAA 122
Qy 155 ArgThrIleGlyTrpAspGlnTyrProHisGlyArgTrpLysAspSerCysAsnProSe 175
Db 123 AAGGACATTAGGTGGGATGAGTATCCCATGACGAGCGGTGGGTGATTCGACATCCATC 182
Qy 175 rTyrGlyAlaLeuSerAspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuVa 195
Db 183 NCATGACACTTACTGTGACACCAAGTTTCAAGACCAAGAGCCGTGTAAAGAGCTCA 242
Qy 195 lGluGluTrpAlaValProLeu 202
Db 243 AGAGAAATGGCTGTTCACATCG 264

RESULT 11
US-09-923-876-476
; Sequence 476, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laljundi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 476
; LENGTH: 265
; TYPE: DNA
; ORGANISM: Zea mays
```



DB 282 DVIEPIKNDAAIRNYGIELAVSLCQELLASGLVPLGHLFHTLNREMATTEVLKRLGMWTE 341  
 QY 120 SKVSSLSLWRRPANVFRVKEVDVRFWAPNPKSYISRTIGWDQYPHGRWGDCSNPSYCAL 179  
 DB 342 DP-RRLPLWALSAPKREEDVRFWASRPKSYIYRQEWDEFFNGRWGNSSSPAFAGEL 400  
 QY 180 SDYQ--FMRPRARDKLVVEAVNPLKSVEDIYERFRLYCIGKLRSN-----PWSELDG 230  
 DB 401 KDYLFLYKLSKSPKEELLKMWGEELTSEASVFEVLYLSGEPNRRGHKVTCLPWN-EP 459  
 QY 231 LOPETKIINEOLEKINTKGFLTINSQPAVNGEKSDPTVWGSGPGGYVQKAYVEFFCSK 290  
 DB 460 LAETSLLKEELLRVNRQGLTINSQPNNGKPSDPTVWGSGPGGYVQKAYLEFFTSR 519  
 QY 291 EKDALVCKKD-RFSLTYMAVNRKDGSKSVNGQTDVNAVTVGVFPFAKEIIPTIVDPVS 349  
 DB 520 ETAEALLQVLLKYLRYNHLVNVKGENITNAPELOPNVAVTVGVFPFAKEIIPTIVDPVS 579  
 QY 350 FNVKDEAFETWRSWASLYPEDEASRKLVEEYGGSHFLVSLVDNDY-INGDLFAVEAD 407  
 DB 580 FMFWKDEAFALWIERWKGKLYEESPSRTIIQYIHDNYFLVNLVNDNDFDNLCLWQVVED 638

RESULT 15  
 AAEL2607  
 ID AAEL2607 standard; Protein; 656 AA.  
 XX  
 AC AAEL2607;  
 XX  
 DT 03-JAN-2002 (first entry)  
 XX  
 DE Human methylenetetrahydrofolate reductase (MTHFR) protein #2.  
 KW Human; methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;  
 KW antisense therapy; EC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;  
 KW pancreas; lung; brain; cystostatic; colon carcinoma; breast; colorectal;  
 KW neuroblastoma; leukaemia.  
 OS Homo sapiens.  
 XX  
 PN US2001025030-A1.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 01-DEC-2000; 2000US-0728910.  
 XX  
 PR 01-MAR-1999; 99US-0258928.  
 XX  
 RA (ROZE/) ROZEN R.  
 RA (SEKH/) SEKHON J.  
 XX  
 PI Rozen R, Sekhon J;  
 XX  
 DR WPI; 2001-638509/73.  
 DR N-PSDB; AAD20463.  
 XX  
 PT New antisense nucleic acids, which are methylenetetrahydrofolate  
 PT reductase inhibitors, useful for treating, stabilizing or preventing  
 PT cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or  
 PT neuroblastoma  
 XX  
 PS Disclosure; Fig 6; 68pp; English.  
 XX  
 CC The invention relates to a non allele-specific antisense nucleic acids,  
 CC which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)  
 CC nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR  
 CC catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to  
 CC 5-methylenetetrahydrofolate, a co-substrate for methylation of  
 CC homocysteine to methionine. The invention provides potential therapy for  
 CC individuals with MTHFR deficiency. The non allele-specific antisense  
 CC nucleic acids are useful for treating, stabilizing or preventing cancer,  
 CC particularly breast carcinoma, colon carcinoma, colorectal carcinoma,  
 CC lung cancer, brain cancer, pancreatic cancer, kidney cell cancer,

CC neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene  
 CC therapy. The present sequence is human methylenetetrahydrofolate  
 CC reductase (MTHFR) protein #2. Human MTHFR gene is mapped to  
 CC chromosome 1p36.3.  
 XX  
 SQ Sequence 656 AA;  
 Query Match 44.8%; Score 979; DB 22; Length 656;  
 Best Local Similarity 44.9%; Pred. No. 6e-89;  
 Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;  
 QY 1 ARVITVQLFYTDITFLFVNDRCQIGITCPIVPGIMPINNYKGFIRMTGFCCKTKIPADIM 60  
 DB 222 ADFIITQFFADFFRFVRACTDMGITCPIVPGIPIQIYQIYHSLQVLKLSKLEVPQEI 281  
 QY 61 AALEPIKNDNEAVKAYGIHLGTEMCKKILAHG-IKTLHLTYLNMKKSALAIIMNLGLIEE 119  
 DB 282 DVIEPIKNDAAIRNYGIELAVSLCQELLASGLVPLGHLFHTLNREMATTEVLKRLGMWTE 341  
 QY 120 SKVSSLSLWRRPANVFRVKEVDVRFWAPNPKSYISRTIGWDQYPHGRWGDCSNPSYCAL 179  
 DB 342 DP-RRLPLWALSAPKREEDVRFWASRPKSYIYRQEWDEFFNGRWGNSSSPAFAGEL 400  
 QY 180 SDYQ--FMRPRARDKLVVEAVNPLKSVEDIYERFRLYCIGKLRSN-----PWSELDG 230  
 DB 401 KDYLFLYKLSKSPKEELLKMWGEELTSEASVFEVLYLSGEPNRRGHKVTCLPWN-EP 459  
 QY 231 LOPETKIINEOLEKINTKGFLTINSQPAVNGEKSDPTVWGSGPGGYVQKAYVEFFCSK 290  
 DB 460 LAETSLLKEELLRVNRQGLTINSQPNNGKPSDPTVWGSGPGGYVQKAYLEFFTSR 519  
 QY 291 EKDALVCKKD-RFSLTYMAVNRKDGSKSVNGQTDVNAVTVGVFPFAKEIIPTIVDPVS 349  
 DB 520 ETAEALLQVLLKYLRYNHLVNVKGENITNAPELOPNVAVTVGVFPFAKEIIPTIVDPVS 579  
 QY 350 FNVKDEAFETWRSWASLYPEDEASRKLVEEYGGSHFLVSLVDNDY-INGDLFAVEAD 407  
 DB 580 FMFWKDEAFALWIERWKGKLYEESPSRTIIQYIHDNYFLVNLVNDNDFDNLCLWQVVED 638

Search completed: February 14, 2003, 23:07:42  
 Job time : 73 secs



```
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 4653
; LENGTH: 159
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID NO. US20010051335A1 700354741H1
; NAME/KEY: unsure
; LOCATION: 14
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4653

Alignment Scores:
Pred. No.: 159
Score: 255.00 Length: 159
Percent Similarity: 94.00% Matches: 43
Best Local Similarity: 86.00% Conservative: 4
Query Match: 11.66% Mismatches: 3
DB: 10 Indels: 0 Gaps: 0

US-09-720-451-6 (1-408) x US-09-294-093B-4653 (1-159)
QY 133 AsnValpHeargValysglAspValargProtlePhePheLeuLysPheValAsnAspCys 152
Db 8 AATGTTTCGATTAAGAGATGTCACCTATATCTGGCCACAGACCAAGAGC 67
QY 153 TyrlleSerArgThrllleGlyTrpAspGlnTrpProHisGlyArgTrpGlyAspSerCys 172
Db 68 TATCTTAAAGACATTAAGTGGATGCGATATCCCATGACGCGTGGGTGATCTCGG 127
QY 173 AsnProSerTyrlGlyAlaLeuSerAspTyr 182
Db 128 AACCCATCATATGAGACACTTACTGACGAC 157

RESULT 15
US-09-734-017A-63
; Sequence 63, Application US/09734017A
; Patent No. US20020142422A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Clippus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides and
; TITLE OF INVENTION: nucleosides
; FILE REFERENCE: BASF-NbE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/771,100
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect
; SEQ ID NO 63
; LENGTH: 409
; TYPE: DNA
```

```
; ORGANISM: Physcomitrella patens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(409)
; OTHER INFORMATION: 62_mm20_c10rev
US-09-734-017A-63

Alignment Scores:
Pred. No.: 409
Score: 213.00 Length: 409
Percent Similarity: 89.80% Matches: 42
Best Local Similarity: 85.71% Conservative: 2
Query Match: 9.74% Mismatches: 4
DB: 10 Indels: 1 Gaps: 0

US-09-720-451-6 (1-408) x US-09-734-017A-63 (1-409)
QY 3 ValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsnAspCys 22
Db 260 GTCATTAATCACCACAGCTGTTTATGATACCATATCTTTTGAATTTGGATGATTC 319
QY 23 ArgGlnIleGlyIleThrcysProIleValProGlyIleMetProIleAsnAsnTyrLys 42
Db 320 CGTCAATTTGGATCAAGGTGCCATTTGTACTGTGATCATGCCCATCAAAATTTACAG 379
QY 43 -GlyPheIleArgMetThrGlyPhe 50
Db 380 GGGCTTTCCTCCGACATGACCACCTT 404
```

Search completed: February 12, 2003, 17:07:24  
Job time : 93 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2003, 23:07:52 ; Search time 3547 Seconds  
(without alignments)  
3347.602 Million cell updates/sec

Title: US-09-720-451-6  
Perfect score: 2187  
Sequence: 1 ARVITQLEFYDTDFLKRNV.....VSLVNDYINGDLFANFADF 408

Scoring table:  
BLOSOM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+P2n.model -DEV-X1h  
-O=/cgn2\_1/USFPO.spool/US05720451/runat\_05022003.074814\_2510/app\_query.fasta\_1.583  
-DB=GenDb1 -GFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=Dits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCMATCH=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US05720451.@cgn2\_1.1616.@runat\_05022003.074814\_2510 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -Delop=6 -Delext=7

Database : GenDb1:  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rnd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hcggo\_hum:\*  
40: em\_hcggo\_mus:\*  
41: em\_hcggo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	1834	83.9	1897	8	AF181967	AF181967 Arabidops
2	1834	83.9	1917	8	AF050434	AF050434 Arabidops
3	1834	83.9	2034	8	AF370515	AF370515 Arabidops
4	1831	83.7	2037	8	AF1245414	AF1245414 Arabidops
5	1777	81.3	1810	8	AF122922	AF122922 Arabidops
6	1777	81.3	1994	8	AF181966	AF181966 Arabidops
7	1777	81.3	2002	8	AF070034	AF070034 Arabidops
8	1762	80.6	2225	8	AF174486	AF174486 zea mays
9	1522	69.6	109741	8	AC004005	AC004005 Arabidops
10	1424	65.1	100906	8	AF124516	AF124516 Arabidops
11	1411	64.5	1785	12	AF441241	AF441241 Synthetic
12	979	44.8	1971	6	AK069359	AK069359 Sequence
13	979	44.8	2187	6	AR144536	AR144536 Sequence
14	979	44.8	2187	6	AR144536	AR144536 Sequence
15	979	44.8	2196	12	HSU09806	U09806 Synthetic c
16	979	44.8	2219	6	A47328	A47328 Sequence 3
17	979	44.8	2219	6	AR097969	AR097969 Sequence
18	979	44.8	2219	6	AR145811	AR145811 Sequence
19	979	44.8	2219	6	AX356174	AX356174 Sequence
20	979	44.8	2219	6	AX452351	AX452351 Sequence
21	979	44.8	2220	6	A47328	A47328 Sequence 1
22	979	44.8	2220	6	AR097968	AR097968 Sequence
23	979	44.8	2220	6	AR145810	AR145810 Sequence
24	979	44.8	2220	6	AX356172	AX356172 Sequence
25	979	44.8	2220	6	HS4237672	AJ237672 Homo sapi
26	895	40.9	37093	8	SPAC56F8	U59849 Caenorhabdi
27	815.5	37.3	38571	3	U39849	U39849 Caenorhabdi
28	812	37.1	81624	3	AC099763	AC099763 Caenorhab
29	778	35.6	3082	8	SCYGL125W	272647 S. cerevisia
30	778	35.6	6812	8	SCSUP44GN	X94106 S. cerevisia
31	629	28.8	3255	8	SPAC343	U57049 Rattus norv
32	629	28.8	42947	8	SPAC343	U57049 Rattus norv
33	578.5	26.5	27120	10	SC036624	AL106929 Mouse DNA
34	452	20.7	1250	10	RNDU57049	U57049 Rattus norv
35	429.5	19.6	198728	8	AL606929	AL606929 Schizosach
36	421.5	19.3	1746	8	D89118	AC025001 Homo sapi
37	415	19.0	159273	2	AC025001	AC025001 Homo sapi
38	404.5	18.5	143039	2	AC125869	AC125869 Rattus no
39	335.5	15.3	181581	2	AC113917	AC113917 Rattus no
40	294	13.4	110000	2	LMFLCHR36.26	Continuation (27 o
41	270	12.3	110000	2	LMFLCHR36.02	Continuation (3 of
42	251.5	11.5	215050	1	AL646057	AL646057 Rattus no
43	250.5	11.5	1376	9	HS4297061	AJ297061 Homo sapi
44	249	11.4	346	11	G71668	G71668 A6499434FM
45	249	11.4	343	11	G70952	G70952 A06717634FM

RESULT 1

ALIGNMENTS

AF181967	1997 bp	mRNA	linear	PLN 13-DEC-1999
LOCUS	Arabidopsis thaliana	methylentetrahydrofolate reductase	MTFR2	
DEFINITION	Arabidopsis thaliana methylentetrahydrofolate reductase	MTFR2		
ACCESSION	AF181967	complete cds.		
VERSION	AF181967.1	GI:5911426		
KEYWORDS				
SOURCE	Arabidopsis thaliana.			
ORGANISM	Arabidopsis thaliana.			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1997)			
AUTHORS	Roje,S., Wang,H., McNeill,S.D., Raymond,R.K., Appling,D.R., Shachar-Hill,Y., Bonnett,H.J. and Hanson,A.D.			
TITLE	Isolation, characterization, and functional expression of cDNAs encoding NADH-dependent methylentetrahydrofolate reductase from higher plants			
JOURNAL	J. Biol. Chem. 274 (51), 36089-36096 (1999)			
MEDLINE	20062814			
PubMed	10593891			
REFERENCE	2 (bases 1 to 1997)			
AUTHORS	Roje,S. and Hanson,A.D.			
TITLE	Direct Submission			
JOURNAL	Submitted (31-AUG-1999) Horticultural Sciences, University of Florida, Hull Road, Gainesville, FL 32611-0690, USA			
FEATURES	Location/Qualifiers			
source	1..1897			
	/organism="Arabidopsis thaliana"			
	/db_xref="taxon:3702"			
	1..1785			
	/EC_number="1.5.1.20"			
	/function="catalyzes the reduction of 5,10-methylenetetrahydrofolate (CH2-THF) to 5-methyltetrahydrofolate (CH3-THF)"			
	/note="NADH-dependent"			
	/codon_start=1			
	/product="methylentetrahydrofolate reductase MTFR2"			
	/protein_id="AAD57788.1"			
	/db_xref="GI:5911427"			
	translation="MKVLDKIQSLADEGKTAFSEFFPERKTEEDGVNLFERRMDRVAAY GTPGADTWAGAGSTADLTDLDIASRMQNVVCVSMNIICTMPEKIDNALETIRSN GIQNVIALRGDRPRGQDFEYVQEGGDFCDALVDNIIRSKYSGDFGTAGYREAHPRVU IGENIALNAENAYOSDLEFKKIKIDAGDADLVNLFQEDTDFLEKFNDDROIGICSPV PGIMEINNYRGFLKMTGFECKTIPVEVMALREIKQNEAVCAVGINGETMCKMLA HGVKILHTLTMERKSALATLIMLGMIDSKISRLSPRRANVRRKEDEVRIETAN RKYSISRTKGMEDEFGSGRSDSASGALSDHOSRPRADKKIOEVRVPLKVE DIOEFKRICGNTKSSPMSIEDGLOPERRITNEDLIYNKSGPITINSQSVAAERS DSPTGMGAGPYGYOKAYILEFCSKEDIKDAYVEKCALPSTTIAVAKGQVNSNA QADVNAVYVSGVPAKELIOTFTYDPAFVNMQDEAFETWSRWANLYPEADPSRNILLE EKNSNITLVSLEVDNIDGIDFVAFVADL"			
BASE COUNT	524 a	362 c	457 g	554 t
ORIGIN				
Alignment Scores:				
Pred. No.:	1,34e-152	Length:	1997	
Score:	1834.00	Matches:	334	
Percent Similarity:	91.15%	Conservative:	37	
Best local Similarity:	82.06%	Mismatches:	36	
Query Match:	83.86%	Indels:	0	
DB:	8	Gaps:	0	
US-09-720-451-6 (1-408) x AF181967 (1-1997)				
QY	1	AlaArgValIleValIThrGlnLeuPheTyrAspThrAspIlePheLeuIysPheValAsn	20	
	111	:::		
Db	559	GCTGACCTCATGTCGTACTGACGCTTTTCATGATCTGATATATCTCCCAAGTTTGAAAT	618	
QY	21	ASPCARGAGGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn	40	
	111			
Db	619	GATTCGCGCAATTCGATTCAGTTGTCCTGATTTCTTCCGAAATTTATATATAC	678	
QY	41	TyrIysGlyPheIleArgMetThrCysIysThrIysIleProIleAspIleMet	60	

Dd	679	TACAGAGGCTTTTGGGATACATCGGTTTGTGCAGACTACAGATACCAAGTTGAGGTGATG	.....	738
Qy	61	AlAlAlLeuGlnPProLLeuYsAspSngLuGuaLAlValLysAlaTyrgLyLeuHsLeu	80	
Dd	739	GCCTCCCTGGAGCGCTATCAGATACAGAGAACGCTGTCAAGCGCTATGATTCACCTT	798	
Qy	81	GLYThrG1WMeCysLyLyLysL1LeuAlaHsGL1LeuYThrLeuHsLeuYThr	100	
Dd	799	GGAACAGAGATGCTAAAAAATCTTGGCTCAGAGATCAAGCTCTTCATCTCTACACA	858	
Qy	101	LeuAsMeG1uLySserLAlaLeuAl1LeuMeAsnLeuGlyLeuL1eGluGluSer	120	
Dd	859	TTGAACTGTGAAGAAATCTGCTTTGCAATATGTATGAATCTGGATGATGATGAGTTC	918	
Qy	121	LysValSerArgSerLeuPProTPArgPArgProAlaAsnValPheArgValLysGluAsp	140	
Dd	919	AAATTTCTCGTTCTTACCTCGAGACGCCCTGCAATATGTTTCCGTAACAGAAAT	978	
Qy	141	ValAlrProL1rPheThrAlaAsnAsArgProLysSerTyrl1SerArgThrL1eGlyTyrr	160	
Dd	979	GTCGGCCCATTTTCTGGGCAAAACCGTCCAAAGAGCTACATTTCTACAAACAAGGCTGG	1038	
Qy	161	AspGlnTyrrProHsGLYArgTyrrGlyAspSerCysAsnProSerTyrgLyAlaLeuSer	180	
Dd	1039	GAAATTTTCCCAAGAGCGGGTGGGGATGCACGAGAGCTTATATGGTGACACTCG	1098	
Qy	181	AspTyrrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGluTyrrAlaVal	200	
Dd	1099	GATCATCACTTCTCACGTCGCCGACGACGTCAGAACAGAAAGCTTCAACAAGAAATGGGTTC	1158	
Qy	201	ProLeuLysSerValGluAspL1LeTyrgLysPheArgPheArgLeuTyrrCysLeuGlyLysLeu	220	
Dd	1159	CCACTGAAAGCTGTGAAGATATTCAAGAGAAATTCAGAGACCTCTCCTGGAACCTT	1218	
Qy	221	ArgSerAsnProTPSerG1uLeuAspGlyLeuGlnProGluThrLysL1eLeuGlu	240	
Dd	1219	AAAAGAGTCCCTGGCTGGAATTATGATGAGACTCCGACGACAGAACAGAAATAAACGAG	1278	
Qy	241	GlnLeuG1uLysL1eAsnThrLysGlyPheLeuThrL1eAsnSerGlnProAlaValAsn	260	
Dd	1279	CAACTGATTAAGTCAACTCCAAAGGCTCTTGACCATCATCATGCAACCATCATGTCAC	1338	
Qy	261	GLYGLuLysSerAspSerProThrValGlyTyrrGlyTyrrProGlyTyrrValTyrgln	280	
Dd	1339	GCGGAGAGATCTGATTCGCCCAAGCTGTGGATGGGAGGTCTGTGGATATGTATACCA	1398	
Qy	281	LysAlrTyrrValGluPhePheCysSerLysGlyLysLeuAspAlaLeuValAspLysCys	300	
Dd	1399	AAGGCTTATCTAGACTTCTTCTGCTCAAGAGAAATATGATGCAAGTGCGGAGAAATGC	1458	
Qy	301	LysAspArgThrSerLeuThrTyrrMetAlaValAsnLysAspGlySerTyrrLysSerAsn	320	
Dd	1459	AAAGCTTTGCCATCATGACTACTACATGCGTGGAACAAAGAGAAAGAGCTGGATGCAAC	1518	
Qy	321	ValGlyLysThrAspValAsnAlaValThrTyrrGlyValPheProAlaLysGlyLeuLe	340	
Dd	1519	ACGCGCCCAAGCTGATGATCAAGCTGTACTGTGGGAGATTTCCCGGCTAAGAAATCATT	1578	
Qy	341	GlnProThrL1LeuValAspProValSerPheAsnValTyrrLysAspGluAlaPheGluLe	360	
Dd	1579	CAACCAACCATTTGTGATCCCGCTACCTTCAACGCTGTGAAGGTGAAGGATTTTAAGCT	1638	
Qy	361	TrpSerArgLyTyrrAlaSerLeuTyrrProGluAspGluAlaSerAlaGlyLysLeuValGlu	380	
Dd	1639	TGGTCAAGAGCTGGGCTACTGTATGTCACCAAGAGCGACCTTCCAGAAACTGCTCGAG	1698	
Qy	381	GluValGlyLysSerHsLAsPheLeuValSerLeuValAspAsnAspTyrrL1eGlnGlyAsp	400	
Dd	1699	GAGGTAAAGACAGCTACTATTGTGTAGCTGTGTGGAGAAAGCATTTACATCATGTGTGAC	1758	
Qy	401	LeuPheAlaValPheAlaAsp	407	
		.....		

Db 1759 ATATGCGGATGTTGCTGAT 1779

RESULT 2

LOCUS AY050434 1917 bp mRNA linear PLN 20-AUG-2001

DEFINITION Arabidopsis thaliana At2g44160/F6E13.29 mRNA, complete cds.

ACCESSION AY050434

VERSION AY050434.1 GI:15215809

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana.

ORGANISM Arabidopsis thaliana.

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1917)

Shinn, P., Chen, H., Cheuk, R., Kim, C.-J., Koesema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J., Liu, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Arabidopsis cDNA clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1917)

Shinn, P., Chen, H., Cheuk, R., Kim, C.-J., Koesema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Direct Submissions

JOURNAL Submitted (02-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

FEATURES

source

1.1917

/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

/chromosome="2"

/clone="RAFL05-16-12(R1336)"

/note="ecotype: Columbia"

1.1405

41.1825

/note="putative methylentetrahydrofolate reductase"

/codon\_start=1

/product="At2g44160/F6E13.29"

/protein\_id="AK91450.1"

/db\_xref="gi:15215810"

/translation="MKYIDKIQSLADGCKIAFSEFFPPKTEGDGNDNFERRDRVAVGPFCDITWAGGSTDADITLISRMKNVVCSSMMHLCTINMPKPERIDHALETITRS

[illegible]

Db 1319 CAACGTATAAAGTCACTCAAGAGCTTGTGACCATCATATACCAACCATCATGATCAAC 1378  
Qy 261 GYGLULYSSERASPERPROTHRALGTYRPGYGLYPROGLYGYTYRVALTYRGIN 280  
Db 1379 GCGGAGAGATCATGATCCCACTGTTGGATGGGAGAGTCTCTGTGATATGTATACCA 1438  
Qy 281 LYSALTYRVALGLUPHERPHECYSSERYSLULYSLEUASPRALALEUVALASPLYSCYS 300  
Db 1439 AAGCTTATCTGAGATCTTCTGCTCAAGAGAAATAGATGAGTGGGAGAAATCTC 1498  
Qy 301 LYSASPARGTHTSERLEUTHRTYMETALAYALASNYASAPGLYSERTPLYSSERA 320  
Db 1499 AAGCTTGTGCATCGATCTTACATGCGCTGGAACAAGAGAACAGTGGTATCGAAC 1558  
Qy 321 VALGTYGLNTHRASPVALASPRALALHTTRPGYALPHEPROALALYSSGUILLE 340  
Db 1559 ACTGCCCAACGCTGATGAACTCTTAACCTTGGGAGCTTTCCCGCTAGAGAAATCAT 1618  
Qy 341 GLNPROTHRTLEVALASPROVALSERPHEASNAVALTRPYLASPGUALAPHEGUIL 360  
Db 1619 CAACCAACCATGTCTCGATCCCGCTTCAACGCTGGAAGATGAGCATTTGAGACT 1678  
Qy 361 TRISERARGTYRPLASERLEUTHRTYPROGLUASPGUALASERARGYSLAUVALG 380  
Db 1679 TGTGTCAAGAAGCTGAGCTTACTGTACCAAGAGCCGCTTCCAGAACTTGTCTGAG 1738  
Qy 381 GLUVALGTYGLYSETHSPHELEUVALSERLEUVALASPNASPLYRILEASNGLYAP 400  
Db 1739 GAGGTGAAGAACGCTACTGATTTGGTGAAGCTTGTGGAGAACGATTCATCATATGTTGAC 1798  
Qy 401 LEUPHEALAYALPHEALASAP 407  
Db 1799 ATATTCGCGGTGTTCTCAT 1819  
RESULT 3  
AF370515  
LOCUS AF370515 2034 bp mRNA linear PIN 30-APR-2001  
DEFINITION Arabidopsis thaliana putative methylesterase/hydrofolate reductase  
(At2g44160; FEEL3.29) mRNA, complete cds.  
ACCESSION AF370515  
VERSION AF370515.1 GI:13877628  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi-  
s (bases 1 to 2034)  
REFERENCE Southwick, A., Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M.,  
Palm, C. J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H.,  
Cheuk, R., Chung, M. K., Hayashizaki, Y., Ishida, J., Kamiya, A.,  
Kawai, J., Kim, C., Lin, J., Liu, S. X., Narusaka, M., Pham, P. K.,  
Sakano, H., Sakurai, T., Satou, M., Seki, M., Shino, P., Yamada, K.,  
Shinozaki, K., Ecker, J., Theologis, A. and Davis, R. W.  
Direct Submission  
Submitted (17-Apr-2001) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
COMMENT e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAPL cDNAs (RAPL cDNA: "RIKEN  
Arabidopsis Full-Length cDNA"): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAPL cDNAs: Southwick, A.,  
Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M., Palm, C. J.,  
Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K.,  
Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shino, P.,  
Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.

Southwick, A. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed  
equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W.  
(SSP/Stanford) contributed equally to this work as PIs.  
location/Qualifiers

FEATURES  
source

1..2034  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/chromosome="11"  
/clone="RAF11-04-D03"  
/note="This clone is in plusescript  
ecotype: Columbia"  
1..2034  
/gene="At2g44160; FEEL3.29"  
91..1875  
/gene="At2g44160; FEEL3.29"  
/cdon\_start=1  
/product="putative methylesterase/hydrofolate reductase"  
/protein\_id="AAK4392.1"  
/db\_xref="gi:13877629"  
/translation="MKVIDKIQSLADEGKTAFSEFFPEKTEGVDNLFERDRMAY  
GPTCDITWAGAGSTADLTLDLSRMQNVVCSMMHLLCTMPEKIDHALLTIRSN  
GIONYLALRGDPHGGODKPYVEGGECDALDVNHRKYGDYFGITVAGYPEAHAPY  
IGENGLASNEAYOSDLEYLKKIDAGADLITQLFYDIDFLKFNOCROGICRPIY  
PGIMPINNYRGELENGECKTIPVVMALRPIKDNPEAVKAGIHGTEPMCKMLA  
HGVSILHYTLNMEKRSALATLMLGWIDSKISRLPMPRRPNRFRTEEDRPTEMAN  
RPKYSISRTKMEDEFPQGRGDSRSXGALSDHOPSPRADRDLOQEMVPLKYE  
DIOERFELCLGKLRKSPWSELDGLDPEKTRIDVETRIEDIKYNSKGLTINSQSVNERS  
DSPVVGWGPVGYTKAVLEFCSKEKIDAVYEKKALPSTITMAVKGQWYNSRA  
QADVAAYWGFPAKEKLIQPTVDPSFVWDEAFETWSRMANLYEADPSRNLL  
EVKNSYLVYSLVENDYINGDIFAVFDL"

BASE COUNT 568 a 386 c 483 g 597 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,46e-152 Length: 2034  
Score: 1834.00 Matches: 324  
Percent Similarity: 91.15% Conservative: 37  
Best Local Similarity: 82.06% Mismatches: 36  
Query Match: 83.86% Indels: 0  
DB: 8 Gaps: 0

US-09-720-451-6 (1-408) x AF370515 (1-2034)

Qy 1 AALARGVALILEVALTHCHLLEUPHERPHEASPRALALPHELEUYSPPHEVALAS 20  
Db 649 GCTGATCTGATGTCTACTGAGCTTCTCATATCTGATATATCTCCATGATTGGAAT 708  
Qy 21 ASPCYARGGLNILEGLYTLERHCYSPROILEVALPROGLYLEMETPROILEASNA 40  
Db 709 GATTGCGGCAATTTGGATGTAGTGTCCCATTTCTTCCGAAATATGCTATATATATAC 768  
Qy 41 TYRGLYGLYPHEILERGWETHRCHLYPHECYLSYSTRHYSILPEPROILAASPILEMET 60  
Db 769 TACAGAGCTTTTGGCGTATGACTGTTTCCCAAGCTAAGATACCAAGTACGATGAGATG 828  
Qy 61 AIALALEUGLUPROILELYSPASPNGLUVALAVALLYALARYGLYLEHISLEU 80  
Db 829 GCTGCTTGAGGCTATCAAGATTAACGAAACGCTGGAAGCGTATGATATTCACCTT 888  
Qy 81 GLYTHGLIMECYLSYSLYSILEUVALHISGLYILEYSTRHLEUNHISLEUTYRTHR 100  
Db 889 GGAACAGAGATGTGTAAGAAAGATGTTGGCTCATGAGATCAAGTCTCTATCATCACA 948  
Qy 101 LEUASMETGLYSSERALEUVALILEUWETASLEUGLYLEUGLUGLUSER 120  
Db 949 TTGACATGAGAAATCTGCTCTTCATATATGATGATGTGATGATGATGATGATGATC 1008  
Qy 121 LYSVALSERARGSERLEUPROTARGARPROLAASNAVALPHEARGVALYSSGUAAP 140  
Db 1009 AAAATTTCTGTTCTTACCCGAGAGAGCCCTGCAATGTTTCCCTACTAAGAGAGAT 1068  
Qy 141 VALARGPROILEPHERTPALASNARGPROLYSSERTYRILESERARGTHRTILEGLYTR 160

Db 1069 GTGGGCCCATTTTCTGGGCAACCGTCCAAAGACTACATTCTAGAACCAAGGCTGG 1128  
Qy 161 AspGlnIleThrProHisGlyThrProGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180  
Db 1129 GAAGATTTCACAAAGCCGGTGGGTGATTCACGACGCTCTTAATGAGTGCACCTCG 1188  
Qy 181 AspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGluTrpAlaVal 200  
Db 1189 GATCATCAGTTCTACGTCGGGAGCAGCTGACAAAGAGCTTCACAAAGAAATGGGTGTC 1248  
Qy 201 ProLeuLysSerValGluAspIleTyrGluArgPheLeuGluTyrCysLeuGlyLysLeu 220  
Db 1249 CCATGTAAGAGTTGAAGATATTCAGGAAATTCAGAGAGAGCTTCGCTGGAAACCTT 1308  
Qy 221 ArgSerAsnProTrpSerGluLeuAspGlyLeuGlnProGluTrpLysIleLeuAsnGlu 240  
Db 1309 AAAAGCAGTCCCTGCTCTGATTAATGATGACCTCCAGCCAGACAGATTAATAACGAG 1368  
Qy 241 GlnLeuGlnLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260  
Db 1369 CAACCTGTAAGAGTCAACTCCAAAGGCTCTGACCATCAATAGCCAACTCATGTCAC 1428  
Qy 261 GlyLysSerAspSerProThrValGlyTyrProGlyLysIleTyrValTyrGln 280  
Db 1429 GCCAGAGATCTGATTCCTCCAACTGTTGATGGGAGGTCCTGTTGATATGATACCA 1488  
Qy 281 LysAlaTyrValGlnPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300  
Db 1489 AAGCTTATCTAGAGTCTCTCTCAAGGAAATATGATGACGCTGGAGAAATGC 1548  
Qy 301 LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrpLysSerAsn 320  
Db 1549 AAAGCTTTCGCAATGATTTCTTACATGCTGTGAGAACAGGAGAACAGTGGATGAC 1608  
Qy 321 ValGlnIleThrAspValAsnAlaValThrTrpGlyValPheProAlaLysGluIleLe 340  
Db 1609 ACTGCCAAGACTGATGATGCTGTAACCTGGGAGGATTTCCCGCTAAGAAATCATTT 1668  
Qy 341 GlnProThrIleValAspProValSerPheAsnValTrpLysAspGluAlaPheGluLe 360  
Db 1669 CAACCAACCAATGTCATGCTCCGCTACCTCAACGTCGAGAGGAGGATTTGAGACT 1728  
Qy 361 TrpSerArgLysTrpAlaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380  
Db 1729 TGGTCAGAAAGCTGGGCTACTGTAACCAAGAGCCGACCTTCAGAAACTGCTGAG 1788  
Qy 381 GluValGlyLysSerHisPheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400  
Db 1789 GAGGTGAAGAACACTACTATTTGGTAACCTGGTGGAGAACGATTACATCAATGCTGAC 1848  
Qy 401 LeuPheAlaValPheAlaAsp 407  
Db 1849 ATATTTCGGGTGTTGCTGAT 1869

RESULT 4  
ATH245414 2037 bp mRNA linear PLN 01-SEP-1999  
LOCUS Arabidopsis thaliana mRNA for methylenetetrahydrofolate reductase.  
ACCESSION AJ245414  
VERSION AJ245414.1 GI:5823582  
KEYWORDS 5,10-methylenetetrahydrofolate reductase; methylenetetrahydrofolate reductase; MTHFR1 gene.  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 2037)  
AUTHORS Ravelle, S., Rebelle, F. and Douce, R.  
TITLE Folate metabolism in higher plants: cloning of a cDNA for 5,10-methylenetetrahydrofolate reductase in Arabidopsis thaliana  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2037)

AUTHORS Ravelle, S.  
TITLE Direct Submission  
JOURNAL Submitted (27-JUL-1999) Ravelle S., DBMS / PCV, URA-CNRS 576, CEA-Grenoble, 17, rue des Martyrs, 38054 Grenoble Cedex 9, FRANCE  
FEATURES  
source 1..2037  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
gene 68..1852  
/gene="MTHFR1"  
CDS 68..1852  
/gene="MTHFR1"  
/EC\_number="1.5.1.20"  
/function="reduction of methylenetetrahydrofolate to methylenetetrahydrofolate"  
/codon\_start=1  
/product="methylenetetrahydrofolate reductase"  
/protein\_id="CA53783.1"  
/db\_xref="GI:5823583"  
/translation="MKVLDKIQSLADEGKTAFFEPFPKTEGVDNLFERMDHVA  
GPTFCIDTWAGGSGTADLFDLIDASRMOSVYCVESMMHLCTNMEPEKIDHALETIRN  
GIOVVALRGDPPHGGDKEVOYEGFDCLDLVNIIRSKYDYGITVAGYEPANHPY  
IGENGASNEAYOSDLEYLKKIDACADLITQLFYDDIDFLKFNDRGICSPY  
PGIMPINNYRGLRMTPGCKTRIPVMAALEIRIDNEAVKAVGHLGTEMCKMA  
HGKSLHLTYLNNKESALILNMLGMDSEKISRELPPRRANPTREKEDVRPLKSYE  
RPSYISRTGMDPFPQGRGDSRSASTGALSDDPSPRRADKKLQDERVYPLKSYE  
DIDKEKEICLGLNLSKSPWSELDGLOPEIRIINEQLIVNSKGLTINSQVNAERS  
DSPVWGSGPVGVYOKAYLEFCSKEKLDAYEKCALKPSITVNAVWKGQWNTV  
QADVNATVAVPFAKEIIOPTIYDPASFPWVDEAFETWSRSMANLYEADPSRNILE  
EVKSYLYLVEINDYINDIDFAVEFDL"

BASE COUNT 583 a 385 c 480 g 589 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,686-152 Length: 2037  
Score: 1831.00 Matches: 334  
Percent Similarity: 91.15% Conservative: 37  
Best Local Similarity: 82.06% Mismatches: 36  
Query Match: 83.72% Indels: 0  
DB: Gaps: 0

US-09-720-451-6 (1-408) x ATH245414 (1-2037)

Qy 1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20  
Db 626 GCTGATCTGATGTTTACACGCTTTCTATGATCATATATATCTCAAGTTGGGAT 685  
Qy 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn 40  
Db 686 GATTGTCGCAAAATGGGATTAAGTTCCTCCATGCTTCGGAATTAATGCTTAATAAC 745  
Qy 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaIleMet 60  
Db 746 TACAGAGGCTTCTTCGATACCTGTTTCGCAAGCTAAGTATACAGTGAAGTGATG 805  
Qy 61 AlaAlaLeuGluProIleLysAspAsnGluAlaValLysAlaTyrGlyIleHisLeu 80  
Db 806 GCTGCTTGGAGCCCTATACAGATACGAAAGAACTGTGAACCTATGATATTCACCTT 865  
Qy 81 GlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThr 100  
Db 866 GGAACAGAGATGTGAAAAGATGTGGCTCATATGATCAAGTCTCTCATCTCAACA 925  
Qy 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSer 120  
Db 926 TTGACATGGAAGAAATTCGCTTCGCAATATGATGATGATGATGATGATGATGATG 985  
Qy 121 LysValSerArgSerLeuProTrpArgArgProAlaAsnValPheArgValLysGluAsp 140  
Db 986 AAAATTTCGCTTCTTACCTCGAGAGCCCTCAAAATGTTTCCGTACTAAGAGAGT 1045  
Qy 141 ValArgProIlePheTrpAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTrp 160

Db 1046 GTGCGCCCATTTCTGGCAAAACCGTCCAAAGACCTACATTCTAGAACCAAGCGCTGG 1105  
 QY 161 ASpgLrYrProHIsIGLrYrPrpGLrYrSpsrYrSasrProserYrGLrYrAlaLeuSer 180  
 Db 1106 GAAGACTTCCACACAGAGCCGGTGGGTGATTCACCAAGTCTTCAATATGGGCACTCTGG 1165  
 QY 181 ASprYrGLrPhemeTarProArGLrAlaArGLrSpsrYrLysLeuValGLrGLrYrPrpAlaVal 200  
 Db 1166 GATCATCAGTTCTCACGCTCCGACAGCTGACACAGAGCTTCAACAGAGTGGGTGTC 1225  
 QY 201 ProLeuYrSerValGLrSpsrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYr 220  
 Db 1226 CCACTGAAAGTCTTCAATATTCAGAGAAATTCAGAGAGCTCTGCTTGGAAACCTT 1285  
 QY 221 ArgSerAspProTrpSerLeuSpsrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYr 240  
 Db 1286 AAAGAGCTCCCTGGTCTAAATTAATGACCCAGCAGAGAGAGATATTAATTAACGAG 1345  
 QY 241 GLrLeuGLrYrLysLeuSerYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYr 260  
 Db 1346 CAAGCTGATAAGTCAACCTCAAGAGCTTTCAGACATCATCAACCACTGATCAAC 1405  
 QY 261 GLrGLrYrSerAspSerProTrpYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLr 280  
 Db 1406 GCCAGAGATCGATTCACCACTGTGGATGGGAGGCTCTGTGGATGATGATACCA 1465  
 QY 281 LysAlaYrYrValGLrPhrPheYrSerYrGLrYrLysLeuSpsrYrGLrYrGLrYrGLrYrGLrYr 300  
 Db 1466 AAGCTTATCTAGATTCCTTCGCTCAAGAGAAATTAATGACAGTGGTGAGAAATTC 1525  
 QY 301 LysAspArGLrSerLeuTrpYrMeAlaValAlaValSpsrYrGLrYrGLrYrGLrYrGLrYr 320  
 Db 1526 AAGCTTGGCCACGATTCATTCATGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1585  
 QY 321 ValGLrYrGLrYr 340  
 Db 1586 ACATGTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1645  
 QY 341 GLrProTrpYrLysLeuSpsrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLr 360  
 Db 1646 CAACCAACCATTCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1705  
 QY 361 TrpSerArGLrYrPrpLysSerLeuYrProGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYr 380  
 Db 1706 TGCTCAAGAGCTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1765  
 QY 381 GLrValGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLrYrGLr 400  
 Db 1766 GAGGTAGAGAGAGCTACTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1825  
 QY 401 LeuPheAlaValPheAlaAsp 407  
 Db 1826 ATATTCGGCGCTTCTTCAT 1846

RESULT 5  
 AY122922 1810 bp mRNA linear PLN 03-JUL-2002  
 LOCUS Arabidopsis thaliana putative methylenetetrahydrofolate reductase  
 DEFINITION MTHFR1 (At3g59970) mRNA, complete cds.  
 ACCESSION AY122922  
 VERSION AY122922.1 GI:21689666  
 KEYWORDS FLU\_CDNAs.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 1810)  
 Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
 Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yu, G.,  
 Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,  
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,  
 Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNT (ORF) clones using the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Kamiya, A., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Narusaka, M., Nguyen, M., Satou, M., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A., Hayashizaki, Y. and Shinozaki, K.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Location/Qualifiers  
 1. 1810  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /clone="018407"  
 /note="This clone is in pUNT 51."  
 ecotype="Columbia"  
 1. 1810  
 /gene="At3g59970"  
 1. 1779  
 /gene="At3g59970"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative methylenetetrahydrofolate reductase MTHFR1"  
 /protein\_id="AA067455.1"  
 /db\_xref="GI:21689667"  
 /translation="MKVYDKIKSVTEGOGAFSPFEPPTKEDEGVENLEFERMDRLVSY GPRPFDITMGAGSTADITLLEIARSMONVCEVETMMLTETNNPIEKIDHALETIRSN GIONVIALGDPHPGQDKFVOEGFACDLVNHISKIDTGFETGAGVPAHPBDV IEADGLATPEESQSDLAIVLKKRVPAGADLVLTQDLETFLEKVNDCQKQIGNCIV PGIMPISNYKFGFLRNAGFCRTKPIAEALALEPIKIDENDEVRKAGVGHFAETMCKKLLA HGITSLHLYTLNVDSKSAIGILMNLGIDESKISLPRPAPRAVVPFTKEDVPIEWAN RPKSYISRTKGWNPDPGRMGDSHSAVSTSDYOPARPGKRGKIOEVPYLNKIE DVOEKFKELCINLKSQSPWSELDIOETKILNOLKINSLEFLINISNPOVYNMAKS DPAIGMGSGPGYVYQKAYLFFPSKSKDLTIYVSKAPFSTIYMAVKNSENVSVTG ESDVNAVTVGVEPAKEVLOPTVDPASFKWKKDEAFIWSWANNLIPBODPSKILLE EVKNSYIVSLVDNNYINGDIFSVFA"

BASE COUNT 508 a 372 c 427 g 503 t  
 ORIGIN  
 3' UTR  
 Alignment Scores:  
 Pred. No.: 1.37e-147 Length: 1810  
 Score: 1777.00 Matches: 319  
 Percent Similarity: 89.41% Conservative: 44  
 Best local Similarity: 78.57% Mismatches: 43  
 Query Match: 81.25% Indels: 0  
 DB: 8 Gaps: 0

US-09-720-451-6 (1-408) x AF122922 (1-1810)

QY 1 AlaATgValIleValIThrGlnLeuPheNyrAspPheLeuPheLeuPheValAsn 20  
 Db 559 GCAGATTTGATGTGACTACAGCTTTCTATGATACATGATATATCCCAAGATTGTGAAAT 618  
 QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn 40  
 Db 619 GATTGCGGCAATCGGATTAATGATGCTTCATGCTTCGAAATATGCTTATTTCAAC 678  
 QY 41 TyrIleGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60  
 Db 679 TACAAAGGGTCTCTGATGAGCTGTCTTGTAAAGCAAGATACCCGCTGACACAT 738  
 QY 61 AlaAlaLeuGluProIleLysAspAsnGluAlaValLysAlaTyrGlyIleHisLeu 80  
 Db 739 GCTGCTTAGAGCTATTAAAGATATATGACGAGCTCTTAAAGCTTATGAAATTCACCTT 798  
 QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThr 100  
 Db 799 GCATCAGAAATGTGCAAAAAGATTTGGCCCATGGAATCATTCCCTTCATCTCTACACA 858  
 QY 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSer 120  
 Db 859 TTGAACGTGACAAATCAAGCTATTGGATTAATGAACCTGTGATGATGAGTCA 918  
 QY 121 LysValSerArgSerLeuProTPrArgProAlaAsnValPheArgValLysGluAsp 140  
 Db 919 AAAATTTCTCGTCTCTACCTTGAGAGCCCTCAAAATGTTTCCGTACTAAGAAAGAT 978  
 QY 141 ValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTTP 160  
 Db 979 GTTCGCCCAATTTTCTGGGCAACCGCTCAAGAGCTACATATGTACAAAGAGGCTGG 1038  
 QY 161 AspGlnTyrProHisGlyArgTPrGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180  
 Db 1039 AATGACTTCCACATGAGAGCTGGGATATTCAACAGTACGATACATACCTTTCG 1098  
 QY 181 AspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGluTPrAlaVal 200  
 Db 1099 GATTATCATTTTGGCGCCCAAAAGAGCTGACAAAGAGCTTACAGCAAGAAATGGCTGTC 1158  
 QY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220  
 Db 1159 CCACATAAAAGCATGAGATGCTTCAGAGAAATTTCAAGAGCTCTCATTTGCAAACTTA 1218  
 QY 221 ArgSerAsnProTPrSerGlnLeuAspGlyLeuGlnProGluThrLysIleIleAsnGlu 240  
 Db 1219 AAAAGAGCCCATGCTGTGAATTAATGATGACTCCAGCCAGAGCAAGATCATTAATGAG 1278  
 QY 241 GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260  
 Db 1279 CAACCTGGAATAATCAACCCCAAGCTTCTGACCAATTAATACCAACATCAGTCAT 1338  
 QY 261 GlyGluLysSerAspSerProThrValGlyTPrGlyGlyProGlyGlyTyrValTyrGln 280  
 Db 1339 GCAGCCAAATCCGATCTCTCAGCTATGTGATGGGTGCTCTGTGTTACGCTACACAG 1398  
 QY 281 LysAlaTyrValGluPhePheCysSerLysGlyLysLeuAspAlaLeuValAspLysCys 300  
 Db 1399 AAGGCTTACTGAGATTCTCTGTTCAAGAGATTAAGACACACTTGAGAGAAATCC 1458  
 QY 301 LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTPrLysSerAsn 320  
 Db 1459 AAGGCTTCTCTCTATCATCATGACATGCGCGGAAACAAATGAGAAATGGGATATCAAC 1518  
 QY 321 ValGlyGlnThrAspValAsnAlaValThrTPrGlyValPheProAlaLysGluIleIle 340  
 Db 1519 ACCGGTAGTCGATGTGATGACAGTACTGTTGGGAGTGTCCAGCTAAGGAGATTATT 1578  
 QY 341 GlnProThrIleValAspProValSerPheAsnValTPrLysAspGluAlaPheGluIle 360

Db 1579 CAACCCAGCATGCTGATTCACACCATTTCAAAAGCTCGAAGAGCAAGCGTTTGAGATT 1638  
 QY 361 TrpSerArgGlyTPrAlaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380  
 Db 1639 TGGTCAAGAGAGCTGGGCTTACTTGTACCCAGAGATGAGCTCTTACAAAGATTGCTCGAG 1698  
 QY 381 GluValGlyLysSerHisPheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400  
 Db 1699 GAGGTGAAGACAGCTACTACTTGTGTAACCTTAGTGGCAACAAATTAATCATATGCGAT 1758  
 QY 401 LeuPheAlaValPheAla 406  
 Db 1759 ATCTCTCTGCTTCTCT 1776  
 RESULT 6  
 AF181966  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1994)  
 Roje,S., Wang,H., McNeill,S.D., Raymond,R.K., Appling,D.R., Shachar-Hill,Y., Bohner,H.J., and Hanson,A.D.  
 Isolation, characterization, and functional expression of cDNAs encoding NADH-dependent methyltetrahydrofolate reductase from higher plants  
 J. Biol. Chem. 274 (51), 36089-36096 (1999)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 20062814  
 10593891  
 REFERENCE  
 2 (bases 1 to 1994)  
 Roje,S. and Hanson,A.D.  
 Direct Submision  
 JOURNAL  
 Submitted (31-AUG-1999) Horticultural Sciences, University of Florida, Hull Road, Gainesville, FL 32611-0690, USA  
 FEATURES  
 source  
 1..1994  
 /organism="Arabidopsis thaliana"  
 /db\_xref="dbEST:W43486"  
 /db\_xref="taxon:3702"  
 63..1841  
 /BC\_number="1.5.1.20"  
 /function="catalyzes the reduction of 5,10-methylenetetrahydrofolate (CH2-THF) to 5-methyltetrahydrofolate (CH3-THF)"  
 /note="NADH-dependent"  
 /codon\_start=1  
 /product="methyltetrahydrofolate reductase MTHFR1"  
 /protein\_id="AAD55787.1"  
 /translation="MKVVDKIKSVTEGQTAFFSEFPPTKTEGVELEFEMDRPLVSYGPPLDIMGAGSTRADLLEIASRMNVICVETMMLCTNNPIKIDHALETIRSNGTQNYLARGPPHGGDQKFOVGGPACALDVNHRSGKGDVETGVAGYEAPDVTEADGIATPESTQSDLATLKKVYDAGADIVTQTFPTDTLFEVNDQOQISINPIYEGITPSNTKGLRMAGCKTKITPAELTAALPEIKDDEAVKATGIEHATECKTLIAHGTSHTYTLVNDKSAIGLIMNLGLIDESTKISLFWRRPAVAFETKEDVPRITWANRPKYSIRTKGWNDPFGHMGWDSHAAYSTLSDVQFARPKGRKYLQDQVVPPLKSIDYQEFKEELCIGNLKSPPWSELDOLEPTYKIMQOLKINSNGFLTINSQVSNVNAKSDPAIGMGGPGGYOKAYLEFCSKDLTIYERSKAFPSITYMAVNSSEWVSNNGESDVNAVYMGVPAPEVIOPTIVDPASPKWKDBAFIWRSMANLYPPDDPSRKLTLEEVKNSYIYSLVNDNYINGDIFSYFA"

BASE COUNT  
 ORIGIN  
 Alignment Scores:  
 Pred. No.:  
 Score:

560 a 404 c 455 g 575 t  
 1..54e-147 Length: 1994  
 1777..00 Matches: 319



```
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAF109-25-L11 (R18407)"
/note="This clone is in a modified pBluescript vector
(FIC-1) as a BamHI/XhoI insert.
ecotype: Columbia"
1..2002
/gene="At3g59970"
5'UTR 1..82
/gene="At3g59970"
CDS 83..1861
/gene="At3g59970"
/codon_start=1
/evidence=experimental
/product="putative methylentetrahydrofolate reductase
MTHFR1"
/protein_id="AA149791.1"
/db_xref="GI:17979047"
/translation="MKVVDKIKSVTEBOGTAPEEFPEPKTEDGVENLPERMDLVSY
GPTFCDITWAGAGSTADLTLEIASRMQVNICVETMMHLCTNMPLEKIDALEIRSN
GIONVALRGDPHGDQKVEYGGFACALDVNHRISKYDYGITVAGYPEAHNPY
IEADGLATPESYOSDLATKKVADADLIYOLEYDIDLEKFNDCROIGINCPY
PGIMPSNKKGEFLMAGFCETKIPALFAALEPIKDNDEAVAYGFEKDEVPKJLA
HGITSILYTLVNDKSAIGILMNLGLIDESKISRLPPRRAPAYFKEDVPRFMAN
RPSYSISRTKGNNDPFGHMDSHSAASTLSDOFAPRRDRDKLQOEWYPLKSE
DVOEKREKICGNIKSSPMSRDLGLOPRTKINDICKINSRGLTINSOPVNAKS
DSPAIKMGPRGGYVOKATLEFECRKDLIDLVKSKAFPSITTMAYNKSNNVSNIG
ESDVNAVTVGVPFAKSVIOPITVDPSKWKMDKAFETWSHMANLYPEDPSKLLLE
EVKNSTYLVLSVLDNNINIGDISVFA"
1862..2002
3'UTR /gene="At3g59970"
BASE COUNT 571 a 409 c 450 g 572 t
ORIGIN
Alignment Scores:
Pred. No.: 1.55e-147 Length: 2002
Score: 1777.00 Matches: 319
Percent Similarity: 89.41% Conservative: 44
Best Local Similarity: 78.57% Mismatches: 43
Query Match: 81.25% Indels: 0
DB: Gaps: 0
US-09-720-451-6 (1-408) x AY070034 (1-2002)
```

```
QY 1 ALaarGValIleValIthrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
111
DB 641 GCAGATTGATTGACTGACTGACTTTTCTATGATGATGATATATCTCCTCAAGTTTGAT 700
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleLeuProIleAsn 40
111
DB 701 GATGTGCGGCAATCGGAGTTAATGTGCCATGTCTCGGAAATTATGCGCTATTTCACAC 760
QY 41 TyrIleGlyPheIleArgMetThrGlyPheCysLysThrLysIleProIleAsnIleMet 60
111
DB 761 TACAAGGGGCTCTGCTATGCGCTGCTTCTGTAAACCAACAGATACCCCTGAGCTCACT 820
QY 61 AlaAlaLeuGlnProIleLysAspAsnGlnGluAlaValLysValArgLysIleHisLeu 80
111
DB 821 GCTGCTTAAAGCCTATTAAAGATTAAGACGAGGCTGTTAAAGCCTATGGAATTCACCTT 880
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThr 100
111
DB 881 GCACAGCAAAATGTCAAAAAGATTGTCGCCATGAAATCACTCCCTTCATCTACACACA 940
QY 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuLeuLeuLysLeu 120
111
DB 941 TTGAACGGGACAAATCAATGCGATTGGGATATATACACCTGGCTGTGATATATAGTCA 1000
QY 121 LysValSerArgSerLeuProIlePargArgProIleAsnValPheArgValLysGluAsp 140
111
DB 1001 AAAATTCTGCTCTACACTTGAGACGCGCTGCAAAATGTTTTCGTAAGAGAGAT 1060
```

```
QY 141 ValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTrp 160
111
DB 1061 GTTCCCAATTTTCTGGGCAACCGTCCAAAGAGCTACTATCTGACAAAGAGGCTCG 1120
QY 161 AspGlnTyrProHisGlyValArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180
111
DB 1121 AATGACTGCCCATGACGACGTTGGGTGATGATCAACAGCTGCACATACAGTACACTTTCG 1180
QY 181 AspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGlnIleThrPheAlaVal 200
111
DB 1181 GATATACAGTTTGGCGCCCAAAAGAGGACAGACAGAGAGGTTCCAGCAAGAGAGGCTCGTC 1240
QY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
111
DB 1241 CCACGTAAGAGATGAGATGATTCAGAGAGAAATCAAAAGAGCTGATGAGAAACTTA 1300
QY 221 ArgSerAsnProTrpSerGluLeuAspGlyLeuGlnProIleThrLysIleLeuAsnGlu 240
111
DB 1301 AAAGAGCGCCATGCTGATTTGATATGATGACTCCAGCCAGACAGACATCAATTAATGAG 1360
QY 241 GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260
111
DB 1361 CAACCTCGAAATATCACTCCAAAGGTTTCTGACCATTAATGACCAACCATCAGTCAAT 1420
QY 261 GlyGluLysSerArgSerProThrValGlyTyrProGlyIleProGlyIleTyrValTyrGln 280
111
DB 1421 GCAAGCAAAATCCGATTCCTCCAGCTATGATGAGGGGTGCTGCTGGTTCACCTACAG 1480
QY 281 LysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
111
DB 1481 AAAGCTTACGAGAGTCTGTTCTGTTCAAAAGATTAAGTTAACACACTTGGGAGAAATCC 1540
QY 301 LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspLysSerTrpLysSerAsn 320
111
DB 1541 AAAGCTTTCTCTTATCACTACATGCGCTGACAAATCAAGATTTGATGATTAAC 1600
QY 321 ValGlyGlnThrAspValAlaValAlaThrTrpGlyValPheProAlaLysGluIleIle 340
111
DB 1601 ACCGGTGAAGCTCGAGTGAATGATGATCTTGGGAGCTGTCCACCTAAGAGGATTAAT 1660
QY 341 GlnProThrIleValAspProValSerPheAsnValTrpLysAspGluAlaPheGluIle 360
111
DB 1661 CAACCGACATGTGTGATCCAGCCAGCTTCAAAATCTCGAAACAGCAAGCGTTTGAGAT 1720
QY 361 TrpSerArgGlyTrpAlaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380
111
DB 1721 TGGTCAAGAGCTGCGCTACCTGTACCCAGAAAGATACCTCTGAAAGTTGCTCGAG 1780
QY 381 GlnValGlyLysThrIlePheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400
111
DB 1781 GAGGTCAAGAAACGCTACTATTGTTAGCTTAAGCTTAAGCAACAATTACATATGATGAT 1840
QY 401 LeuPheAlaValPheAla 406
111
DB 1841 AATCTCTGCTGCTTGTCT 1858
111
RESULT 8
LOCUS AF174486 2225 bp mRNA linear PLN 13-DEC-1999
DEFINITION Zea mays methylentetrahydrofolate reductase mRNA, complete cds.
ACCESSION AF174486
VERSION AF174486.1 GI:5802605
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2225)
REFERENCE
Rojas, S., Wang, H., McNeill, S.D., Raymond, R.K., Applin, D.R.,
Shachar-Hill, Y., Boehrner, H.J. and Hanson, A.D.
Isolation, characterization, and functional expression of cDNAs
encoding NADH-dependent methylentetrahydrofolate reductase from
```

higher plants  
JOURNAL J. Biol. Chem. 274 (51), 36089-36096 (1999)  
MEDLINE 20062814  
PUBMED 10593891  
REFERENCE 2 (bases 1 to 2225)  
AUTHORS Wang, H. and Boehrert, H.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-1999) Department of Plant Sciences, University of  
Arizona, Forbes Building, Room 303, Tucson, AZ 85721, USA  
FEATURES  
source  
1..2225  
location/Qualifiers  
/organism="Zea mays"  
/db\_xref="caxon:4577"  
77..1858  
/EC\_number="1.5.1.20"  
/function="catalyzes the reduction of  
5,10-methylenetetrahydrofolate (CH2-THF) to  
5-methyltetrahydrofolate (CH3-THF)"  
/note="NADH-dependent; MTHFR"  
/codon\_start=1  
/product="methylene tetrahydrofolate reductase"  
/protein\_id="AAD51733.1"  
/db\_xref="GI:5802606"  
/translational\_table="MKVTEKILAEAGDRTAFSPPEPPKTEEGVENLEPERMDRAVAH  
GPSFCDITMGAGSTADLTLEIANRMGMNVCFTMMHLCCTNMPEVKIKIDIALETKSN  
GIONVALRCDPBPBGCKRFEVCGSCALCDIVQIIRAKYGFPGTIVAGYEPKAPDA  
IOGEGATLEAYSNDAIATKRVNAGADLYDTFDIDTLEFVNDCKQITCPYV  
PGIMPINNYKEFLNRGTCKTIPSETALADLPKIDEEAVRQYGIHDETCKKILA  
TGKTLHLTYLMDKSAIGILMNGLIEESKVSRLPWRPATVPRVKEDVPIRMAN  
RKSKYLKRLTGMDOYPHGMWDSRNPSTGALTDHOFTRPGRGRKLOEEMAPLKSVE  
DISERTNEFCGKLTSSWSELIDGLQPEYKIIDQIVINIKGFTLINSQPAVLSKE  
DSPTVGMGPGGYVOKAYLEFFCAREKLDOLIEKIAKPSLYLIYNVDGETFNIS  
PAAVNAVTVGVEPKEIIOPTVVDHVASMKDEAFETMRGMGCMFPGSDSRELE  
KVQKTYIVSLVDNDYVGGDLFAAFK"

BASE COUNT 576 a 489 c 573 g 587 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,73e-146 Length: 2225  
Score: 1762.00 Matches: 324  
Percent Similarity: 88.64% Conservative: 35  
Best Local Similarity: 80.00% Mismatches: 46  
Query Match: 80.57% Indels: 0  
DB: Gaps: 0

US-09-720-451-6 (1-408) x AF174486 (1-2225)

QY 1 AAlaArgValIleValThrcInleupheryrAspThrAspIlePheLeuLysPheValAsn 20  
||| :::  
Db 635 GCTGACCTTATGTGTACACAACTTTCTATGATACCGACATCTTCTCAAGTTTGTGAAT 694  
QY 21 ASpCYArGslInleGlyIleThrcysProIleValProGlyIleMetProIleAsn 40  
||| :::  
Db 695 GACTGCCGCCAAATGTGATTTACTTGCCATGTTCTGCGCAATGACCAATAAATAC 754  
QY 41 TyrlYsGlyPheIleArgMetThrcIlyPheCysLysThrlYsIleProAlaAspIleMet 60  
||| :::  
Db 755 TACAAGAGTTCTCTCGCATGACGCGGTCTGCACAAATAGATACCTTCAGACATCACT 814  
QY 61 AAlaIleuGlnProIleIlySaspAsnGluAlaValLysAlaIlyrGlyIleHisLeu 80  
||| :::  
Db 815 GCTGACCTGATCTATCAAGACAAATGGAGGCTGTTAACAAATATGGAATCCACCTT 874  
QY 81 GLyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrlleHisLeuYrThr 100  
||| :::  
Db 875 GGAACCTGAGATGTGCAAGAAATTTCTGTACTGCGCATTAAGACTTTCACCTTTACACA 934  
QY 101 LeuAsnMetGluYsSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSer 120  
||| :::  
Db 935 CTAAACCTGGACAAAGTCTCTATAGCAATTTGATGAATCTTGATTAATTGAGAGACTCC 994  
QY 121 LysValSerArGserLeuProItrParGArProAlaAsnValPheArgValLysGluAsp 140  
||| :::

Db 995 AAGGTTTCAGGCGCATTTACCTTGAGCGCCAGGAGCTAAATGTTTCCGTTAAAGAGAT 1054  
QY 141 ValArgProIlePheThrPalAsnArgProLysSerTyrIleSerArgThrlleGlyTyr 160  
||| :::  
Db 1055 GTTCCACCTTATTTCTGGGCAACAGACCAAGACCTATCTTAAAGACATTTAGGTGG 1114  
QY 161 ASpGlnTyrProHisGlyArgTyrGlyIleAspSerCysAsnProSerTyrGlyAlaLeuSer 180  
||| :::  
Db 1115 GATCAGATATCCCATGGAGCGGTGGGTGATTTCTCGGAACCCATCATATAGGAGCTTACT 1174  
QY 181 ASpTyrGlnPheMetArgProAlaArgAspLysLysLeuValGluGluTrrAlaVal 200  
||| :::  
Db 1175 GACACCAAGTTCACAGACACGAGCGGTGTGAAGACCTCCAGAGCAATGGCGCTT 1234  
QY 201 ProLeuYsSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220  
||| :::  
Db 1235 CCATGTGAATCTGTGGAGGAGCATTAAGAGCTTACAAACTTGTGCAAGGAAATCTC 1294  
QY 221 ArgSerAsnProItrPserGluLeuAspGlyLeuGlnProGluThrlYsIleIleAsnGlu 240  
||| :::  
Db 1295 ACAAGCAGCCCATGCTCAGATTGACGGCTCTTCAACAGACAAAGATTATGATGAC 1354  
QY 241 GlnLeuGlnLysIleAsnThrlYsGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260  
||| :::  
Db 1355 CAGTTGATGAATTTATACCAAGAGGTTCTTACATTAACAGCAACCTGCTGTAAT 1414  
QY 261 GlyLysLysSerAspSerProThrValGlyTyrGlyGlyProGlyGlyTyrValTyrGln 280  
||| :::  
Db 1415 GGAAGAAATCCACACCTCGCTACTGTGGTGGGTGGTCCGAGGCTTACGTTATACAG 1474  
QY 281 LysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300  
||| :::  
Db 1475 AAGGCTTACCTCGAATTTCTTCTGCGCAAGAGAAAGTGGACCAACTATGAGAAATC 1534  
QY 301 LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrrLysSerAsn 320  
||| :::  
Db 1535 AAGCATTTCCCTCTCTCACTTACATTTGCTGGAACAAAGATGAGAAACATTTCCAAAT 1594  
QY 321 ValGlyGlnThrAspAlaAsnAlaValThrTrrGlyValPheProAlaLysGluIleIle 340  
||| :::  
Db 1595 ATTTCACCGAAGCGCGGATGATGCTGACGCTGGGCTTCCCTGCGCAAGAAATATATC 1654  
QY 341 GlnProThrIleValAspProValSerPheAsnValTrrLysAspGluAlaPheGluIle 360  
||| :::  
Db 1655 CAGCTACGGGTGGGTGATCAGTCACAAAGTTTATGTTTGAAGAGCAAGCAATTTGAATC 1714  
QY 361 TrpSerArgIlyTrrPalSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380  
||| :::  
Db 1715 TGGACTCGGGGGTGGGTTGCATGTTCCCTGAGGGTGAATTCAGAGGAGCTACTAGAG 1774  
QY 381 GluValGlyLysSerHisPheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400  
||| :::  
Db 1775 AAGGTTCAAGAAAGCTTACCTGTGATGACCTGTGACAAAGACTTACGTCAGGGGAG 1834  
QY 401 LeuPheAlaValPhe 405  
||| :::  
Db 1835 CTGTTGCTGCTTC 1849  
||| :::  
RESULT 9  
AC004005  
LOCUS 109741 bp DNA linear PLN 11-MAR-2002  
DEFINITION Arabidopsis thaliana chromosome 2 clone FBE13 map C1C10F02,  
C1C02E07, complete sequence.  
ACCESSION AC004005  
VERSION AC004005.3 GI:20197067  
KEYWORDS HMG.  
SOURCE  
ORGANISM Arabidopsis thaliana.  
Arabidopsis thaliana.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 109741)  
AUTHORS Rounsley,S.D., Kaul,S., Lln,X., Ketchum,K.A., Crosby,M.L.,

Brandon, R. C., Sykes, S. M., Mason, T. M., Kerlavage, A. R., Adams, M. D., Somerville, C. R. and Venter, J. C.  
Unpublished  
2 (bases 1 to 109741)  
Lin, X.  
Direct Submission  
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
3 (bases 1 to 109741)  
Town, C. D. and Kaul, S.  
Direct Submission  
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, cdtown@igf.org  
On Apr 18, 2002 this sequence version replaced gi:6598402.  
FEATURES  
source  
1..109741  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/map="C1C10P02, C1C02E07"  
/clone="F6E13"  
complement(11..4149)  
/note="overlap with BAC clone F18019 (AC002333:1..4149)."  
repeat\_region  
complement(1191..3221)  
/rpt\_family="AT\_rich"  
3302..4750  
/gene="At2g43880"  
/note="F6E13.1"  
join(<3302..3748,3846..4135,4219..4447,4532..>4750)  
mRNA  
join(<3302..3748,3846..4135,4219..4447,4532..4750)  
join(3302..3748,3846..4135,4219..4447,4532..4750)  
/gene="At2g43880"  
/codon\_start=1  
/product="putative polygalacturonase"  
repeat\_region  
complement(14748..17292)  
/protein\_id="AAC23426.1"  
/db\_xref="GI:3212875"  
/translation="MANNISFPCALITFEFINKSHAMPFNVOYRGARQDRA  
DARKSLTAMSLACGSRARAMYVRGTYLVKNLTFWGCCKNIITFEKNGTLVAVAN  
WDINGNGWILPAKYNRTSYGCTIDARAGWSCRKSGHCGARASISPEWCNVL  
LGSLSFSQNNHVTYVHSSNRIENVRIRASGSPNIDGTHVOSSGVTTSISGCTIGAT  
GDICIALFSQSNRIWLERVNGSPGHGISIGSLGDYANEGVONVTYTSVFTKTNGV  
RIKTWARSPRGVNNVFNRLIMNNVENPVIIDQNYCPNGKCPROSSVKLSGYTFA  
NIKGTSTPIAMKIDCGSNHCTGLRLDIKLTVMRSSASYCRNAHGASGVMPRN  
CM"  
repeat\_region  
complement(5593..5613)  
/rpt\_family="AT\_rich"  
repeat\_region  
complement(6730..6753)  
/rpt\_family="AT\_rich"  
7192..8658  
/gene="At2g43890"  
/note="F6E13.2"  
join(<7192..7632,7743..8032,8122..8350,8440..>8658)  
mRNA  
/gene="At2g43890"  
join(7192..7632,7743..8032,8122..8350,8440..8658)  
/gene="At2g43890"  
/codon\_start=1  
/product="putative polygalacturonase"  
repeat\_region  
complement(16570..16608)  
/protein\_id="AAC23398.1"  
/db\_xref="GI:3212847"  
/translation="MDNNKLIIVAVLIMFESSFLMKSSSTASNVVVSFGAKPDGRDTS  
TKAFLGAMQACRSAAVTVYRGSLFLKPYEFGPCRSRTIFQYGTIVAPSDYRG  
LGNNGWILFVKYNRISITIGTLIDARAGFWCRKSGKSPYAGSMFPMAMANDVVS  
GLTINSQTHLYVINSNNVYAKYKLVAPDOSPNIDGTHVOSSGVTTSISGCTIGAT  
DCISIGPGRNLTMSKLNCGPHGISIGSLGRDANAGYENTTLINSVSGSDNGYRI  
KTMAROSTGVNRNVLEQNLIMKNVQNPILVDQNYCPNSGCPQSGSVKLSQVYKNI  
QGTSTFOALTFDCSRSNPCOALRLHDIKLTFRNGRSTATCKNIKGVKAGVMPDQGL  
"  
gene  
complement(9101..14090)  
/gene="At2g43900"  
/note="F6E13.3"  
complement(join(<9101..9952,10049..10409,10500..10623,  
10717..10984,11093..11133,11219..11384,11587..11671,  
10717..10984,11093..11133,11219..11384,11587..11671,

CDS  
11768..11895,12004..12811,12910..13162,13259..>14090))  
/gene="At2g43900"  
complement(join(9101..9952,10049..10409,10500..10623,  
10717..10984,11093..11133,11219..11384,11587..11671,  
11768..11895,12004..12811,12910..13162,13259..14090))  
/gene="At2g43900"  
/codon\_start=1  
/product="putative inositol polyphosphate 5'-phosphatase"  
repeat\_region  
complement(15108..15141)  
/rpt\_family="AT\_rich"  
repeat\_region  
complement(16570..16608)  
/rpt\_family="AT\_rich"  
repeat\_region  
complement(17642..17687)  
/rpt\_family="AT\_rich"  
repeat\_region  
complement(17744..17794)  
/rpt\_family="AT\_rich"  
repeat\_region  
complement(17926..17970)  
/rpt\_family="AT\_rich"  
repeat\_region  
complement(18152..18196)  
/rpt\_family="AT\_rich"  
gene  
complement(19342..21515)  
/gene="At2g43920"  
/note="F6E13.5"  
complement(join(<19342..19624,19773..19851,19988..20034,  
20130..20229,20308..20408,20521..20607,20828..20958,



Db 94670 CAAGGAGCTGCTGGTGAACCTTAAGACAGTCCCTGCTGATTAAGATGACTCCA 94729  
 QY 232 nProgluThrIysIleIleAsnGluInLeuGluIlysIleAsnThrIysGlyPheLeuTh 252  
 Db 94730 GCCAAGACAAAGGATAAATAACGACCACTATATAAGTCAACGCCAAGGCTTCGTGAC 94789  
 QY 252 rIleAsnSerGlnProAlaValAsnGlyIuIysSerSerPheProThrVal ----- 269  
 Db 94790 CATCAATAGCCACACATACAGTCAACGCCAGAGATCTATTCGCCAATGCTTGGTGAGCA 94849  
 QY 269 ----- 269  
 Db 94850 ATCAGCTCGTATATCTCTGTTCTCATAGCAACAGAGATTTTGATGCTGATTGATGCA 94909  
 QY 270 ----- GlyTPDgLygLyProGlygLyTyrrAlTyrrGlnLy 281  
 Db 94910 TACTTGCTGCTGTATGATTTGACGAGATGGGAGGCTCTGTTGATGTATACCAAA 94969  
 QY 281 sAlaTyrrAlGluPhePheCysSerIysGlyuIysLeuAspAlaLeuValAspIysCysAlY 301  
 Db 94970 GGCTTATCTAGAGTCTCTGCTCAAGAGAGAAATTAATGACAGTGTGAGAAATGCCAA 95029  
 QY 301 sAspArgThrSerLeuThrTyrrMetAlaValAsnIysAspGlySerTyrrIysSerAsnVa 321  
 Db 95030 ACCTTGCCATGATTAATCTACATGCTGTGACAAAGAGACAGTGGTATGACACAC 95089  
 QY 321 lGlyGlnThrAspValAsnAlaValThrrTyrrGlyAlaPheProAlaIysGlyIleIleG1 341  
 Db 95090 TSCCCAAAGCTGATGTAAGTGAAGCTGTAAGTGGGAGCTTTCCCGCTAAGAAATCATTCA 95149  
 QY 341 nProThrIleValAspProValSerPheAsnValTyrrIysAspGlyuAlaPheGluIleTr 361  
 Db 95150 ACCAACCATGTCGCATCCCGCTACAGCTTCGGAAGATGAGACATTTGAGACTTG 95209  
 QY 361 pSerArgGlyTyrrPalasSerLeuTyrrProGluAspGluAlaSerArgIysLeuValGluG1 381  
 Db 95210 GTCAAGAACCTGGGCTAACTGTACCCAGAACCGACCTTCAGAAACTGCTCGAGGA 95269  
 QY 381 uValGly ----- 383  
 Db 95270 GGTAAG-TTCTATATTCTTCTGATGATCATATTGTTGTTTGGACAAATGCTTT 95328  
 QY 384 ----- GlySerHisPheLe 388  
 Db 95329 ATTATAAAGACTAATGACATTTTGATCCAAATAATGACAGTGAGAACAGCTACTATT 95388  
 QY 388 uValSerLeuValAspAsnAspTyrrIleAsnGlyAspLeuPheAlaValPheAlaAsp 407  
 Db 95389 GGTAACTTGTGGAGAACGATTAATCATATGATGATATTCGGGGTGTTCGTGAT 95446  
 RESULT 10  
 ATP24G16 100906 bp DNA linear PLN 21-FEB-2000  
 LOCUS Arabidopsis thaliana DNA chromosome 3, BAC clone F24G16.  
 DEFINITION Arabidopsis thaliana DNA chromosome 3, BAC clone F24G16.  
 ACCESSION AL138647  
 VERSION AL138647.1 GI:6899904  
 KEYWORDS  
 ORGANISM Arabidopsis thaliana.  
 Arabidopsis thaliana.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 100906) Modesto,D., Pigazzi,M., Valle,G.,  
 D'Angelo,M., Vezzi,A., Mayer,K.F.X., Quetier,F. and Salanoubat,M.  
 AUTHORS Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F. and Salanoubat,M.  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 100906)  
 AUTHORS EU Arabidopsis sequencing project.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-FEB-2000) MIPS, at the Max-Planck-Institut fuer  
 Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:  
 lemcke@mps.biochem.mpg.de, mayer@mps.biochem.mpg.de, project  
 Coordinator: Marcel Salanoubat and Francis Quetier, Groupement

COMMENT  
 d'Interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue  
 Gaston Cremlieux, BP191, 91006 Evry Cedex, France;  
 http://www.genoscope.cns.fr  
 Information on performance of analysis and a more detailed  
 annotation of this entry and other sequences of chromosomes 3, 4  
 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..100906  
 /organism="Arabidopsis thaliana"  
 /variety="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 1..1731  
 /note="overlap to BAC T16L24, please refer EMBL Acc. no  
 AL138659 for analysis and annotation"  
 2..1888  
 /gene="F24G16.10"  
 complement(42..1888)  
 /note="Contains Protein kinases signatures and profile  
 AA339-362; Protein kinases signatures and profile  
 AA458-470"  
 /codon\_start=1  
 /product="receptor lectin kinase 3"  
 /protein\_id="CAB75793.1"  
 /db\_xref="GI:7019668"  
 /translation="MSHKVLOIVLVLLTFSTHNSNGNFMEEAANGCYCLT  
 NTKHSYGAPNNTPPYIKNSSRSENIIFGIVPEHKQSGHGMARFSPTRLPASP  
 DOYLGFENENKSKASNVYATIEDLRKEEGEDIDNHYGNTNGISVASAGY  
 DDDGDFKRLSLISTVWRSLVSYSTEDQLNVTLPATISVPKPSLSLRDLSPI  
 FLEETVIGFTASIGALYVWQFSEEGVYIPMEDLGIVPTLPYPRKSKYDRRI  
 LAWLCLTAVTALVASIGEFVEFVRKRYKEVLEMEIONGHRSYKELFNATGFK  
 EKOLGKGGGVYKGMIPGSDAEIVKRTSDSRGMSFPAEISTIGRLHPMLVR  
 LGYCKRKENTLYVDFPMNGSLDRCTFSNNENRERLTMRORRRIKDVATALLH  
 HQMVQVITYHRDIKRPANVLLDHGMARLDRLGATKYIDGQEPQSRVAGTGLTAPL  
 LKTRGATSTDYVAGLVMEVCGRLIERPAENAEVLDWMLELSEKLFDAE  
 ESIRROENKGEIEVLKLGILCAHRETLIRPMASVLIQLNGVSHLPNNLDVVAE  
 RLKGI"  
 complement(2..1888)  
 /number=1  
 2685..4651  
 /gene="F24G16.20"  
 complement(join(2685..4127,4214..4651))  
 /note="strong similarity to receptor lectin kinase 3 -  
 Arabidopsis thaliana, EMBL:U93161; Contains Protein  
 kinases signatures and profile AA309-332; Protein kinases  
 signatures and profile AA425-437"  
 /codon\_start=1  
 /product="receptor lectin kinase-like protein"  
 /protein\_id="CAB75794.1"  
 /db\_xref="GI:7019669"  
 /translation="MPSLKVLIHVLLVLTSSSTYNSNGNMWLTGSAADNSIGPTI  
 LNTWKSGCGCTFNMSIPIKDSSEFPHLFGIVPEHKQSGHGMARFSPTRLPASP  
 SDOYIGLEFNEITNGKSNHYATIEDLRKEEGEDIDNHYAMVRSLVSYSHDQ  
 LNTLPLPAETLPVPRKPLSLNDSIPYLEMEYGYTASSTIGAFHMSVATPK  
 VENPTEWEIFVPTLPYPRKSDRTRKILAVCLTAVFAVYASGICFVFRHKKVK  
 EVLEMEIQGPRFAVKELNATKDKFKGKLGKGFQVFGTLPGSNATVAKRT  
 SHDSRQMSSEFLAISTIGRLHPNVLRLGCRKENTLYLVDFTPGSLKYDRN  
 ENOERLTMRORRRIKDVASALLHQQEWQVLIHHDIDPANVLLDHENARITGEG  
 AKTYDGDQRTSRVAGTGYINAPELLRTGRATSTDYVAFGLVMEVCGRMLER  
 APNEEVLDWMLEWSEKLFPAEESIRROENRETELKLGILCAHRETLIRP  
 MASVNIILNVSQLPNNLDVVAENLRMPETSIIEVLGLNLYSGVMTLSPFLSH  
 GR"  
 complement(2685..4127)  
 /number=1  
 complement(4128..4213)  
 /number=1  
 complement(4214..4651)  
 /number=2  
 5498..8175  
 /gene="F24G16.30"  
 complement(join(5498..5626,5840..5920,6042..6101,  
 6253..6332,6432..6483,6611..6748,7011..7274,7405..7462,  
 7549..7658,7819..8175))

/note="several differences to EMBL:X81973: Contains  
ATP/GTP-binding site motif A (P-loop) AA169-176; Cysteine  
synthase/cystathionine beta-synthase P-phosphate  
attachment site AA143-161"  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="CA875795.1"  
/db\_xref="GI:7019670"  
/translation="MYAMIMASRFNREAKLASQILSTLLGNRSCTSMATSSALL  
NPLSSSSSTLRPRPCSPETISLSSTSSSDSLAKKRSRFPAGSEBDEPSVCEAY  
KRETCGPGDNLADNWSQLGKTPMYLINSIANGVANIATLEIMEPCSVKDRIGYS  
MYTDAQKGFISPGKSVLEPTSGNTGILAFIASSRGYRLITLMPASSMERRVLLK  
AFGAELVLDKPSKMGAVOKAEELIKNTPDAYMLQEDNPANPKIHETTGPETMD  
TKGKVDIFVAGTGGTITGVGRFIEKKNPTQVIGVEPTESDILSGGKPGPHKLOGI  
GAGFIPKNDOKIMDEVIAISSEAEITEKOLAKELGKLWVGTSGAAAAAKVAKRP  
ENAGKLIATVNSMNAFQVFPSPGEXELSTPLPQSIREEVEMQPEV"  
/number=1  
/complement(5627. .5839)  
/number=1  
/complement(5840. .5920)  
/number=2  
/complement(5921. .6041)  
/number=2  
/complement(6042. .6101)  
/number=3  
/complement(6102. .6252)  
/number=3  
/complement(6253. .6332)  
/number=4  
/complement(6333. .6431)  
/number=4  
/complement(6432. .6483)  
/number=5  
/complement(6484. .6610)  
/number=5  
/complement(6611. .6748)  
/number=6  
/complement(6749. .7010)  
/number=6  
/complement(7011. .7274)  
/number=7  
/complement(7275. .7404)  
/number=7  
/complement(7405. .7462)  
/number=8  
/complement(7463. .7548)  
/number=8  
/complement(7549. .7658)  
/number=9  
/complement(7659. .7818)  
/number=9  
/complement(7819. .8175)  
/number=10  
/gene="F24616.40"  
/complement(join(12111. .12392,12490. .12723,12815. .13255,  
13361. .13588,13679. .14482,14572. .15081,15220. .15462,  
15581. .15751,16213. .17432,17543. .17675,17797. .18392,  
18486. .18504))  
/note="contains ATP/GTP-binding site motif A (P-loop)  
AA518-525; WW/rsp5/WWP domain signature and profile  
AA514-540"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CA875796.1"  
/db\_xref="GI:7019671"  
/translation="WDLHPGGSKRTSVVVTLDTEGVYIASLLSKADQVIYIDPT  
TGLIRYNGKPGIDNFKSEREALDYITNGSRGVSSVYARAIIIGYVIGSGMLVAT  
RLNPSIPDLPGGCVYVAESQWKIPLYNPQPGKGTNIOELTDLIDGKHYFD  
TRDIRPPSRMPLQSPDDEFVNNRWLSVPRKNTGLDEHCYIILQGFARPPSSSQ  
LEGVIALMARSRRLHPTRTYLRKINSCTGNEVECDLWIKRNGQSIATFNSYIW  
RGTPTIWWGAEIKMTAAEIVYADRDYKSGSTYYQRLSKRDYTRNLDAFVENOK

KKAFVPIVCVNNLLSGSGKSCITLVQHFEESSMNRKSSGKLPTRVHLIYNDHNASV  
LKGRQITBGLMNLKSPYATIGISEDDYPSRRLKDCGEVYICDIDIGACFLRSH  
QNGVIRENCADSLORTNAASFPGLOVFEVRCRLGISLTDIGYGNHNVNNGGYNA  
PLPQWKEKRADAVYKGSYIDHNTKTTWSPHCPDKPMKRLDMFEFKSTILSPYS  
ELADLFLOQGDIFHATLYTGSKAMHSIYNFPEESGAFKQFSAOKMKITLQRRYKN  
AMVSSROKOLEMFLGMRLEFKHPSIYVOPHLICPOAADIVELFYTESPCHCOLL  
TISHGADDTLQPSVYDRTGTHIEDLKIYVNGASIPACACTNLVPLRPISEDA  
VTGNGARLHEHDTSSLSLIDFEELBEQDLFLRVAVTTPYAGAVRIPMTLQILEV  
GTSLEPMGFCEKRTGRLALDELARKPDEDSPESSCDLNPFAATYSLOAEVSTPQ  
QKDPFNSLIDLNGEDSSDPFQPVVEGIASGNDMLDLDEAVVEYSGDTPVPG  
SVPNKRPKDSGAHLVNLCKSLGPMVKNYFPVATVCCSDCMSLAKLFEVE  
AMKLEIRLRNLINSAERDRLSLIGDPTINNSYDELTYGRCLKINALAVNGO  
ASLDDITIASIGLEKLENNVIDENITRIGEGCGMCOVRAVNSPVGSSSRG  
ESSGVLPFOCMKACRCKCAGKALLSKSYSDPTANGGSLADVSATSGDHYIC  
KKCCSVLEBLIYDYRVNMSGRVNDAGREALENEVGSNTNHLAVRQSPN  
REDNFNLQILGKEESISEPFSFLKRVETADSAFESLLPLPLINASSNAWKAAP  
SADSEVAEIVLNTLSDVSVILVSPGCVDAAPVQVIAASDINKERTLNGKNDY  
QSFYRSSPELSGSESGRABRIKFAKNVRCRIITLRLRGLSSSVLDKNIN  
LISDENFAPIPRRASFGATIENDPCIHAKHILINVTARDKTLQESMSVND  
RAPRLNRLIPLDTEREMENDIVLEIYLOASPLAAGFLDASATKPRATHSPSSD  
VDIWDPTSVIMEDRHVSFALITYIOVSLOQYKVTIATRYLPEARGLKUIYDFPQ  
IQACRVSKLGDVAAFTDPAEAEVADLSRASPAPAGLSIANRIKLYYADPEVGKW  
TSLSSV"

exon  
/number=1  
/complement(12111. .12392)  
/number=1  
/complement(12393. .12489)  
intron  
/number=1  
/complement(12490. .12723)  
exon  
/number=2  
/complement(12724. .12814)  
/number=2  
/complement(12815. .13255)  
intron  
/number=3

Alignment Scores:  
Pred. No.: 2,956-114 Length: 100906  
Score: 1424.00 Matches: 320  
Percent Similarity: 47.51% Conservative: 42  
Best Local Similarity: 41.99% Mismatches: 43  
Query Match: 65.11% Indels: 360  
DB: 8 Gaps: 8

US-09-720-451-6 (1-408) x ATF24G16 (1-100906)

QY 1 ALaarGValIleValIThGlnLeuPhenTyraSPHraSPHleuLysPheValAsn 20  
||| ::  
Db 84877 GCAGATTGATTGCACACGCTTTCTATGATACGTATATATTCCTCAAGTTGTGAAT 84936  
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40  
||||| ::  
Db 84937 GATTGTGCGCAAAATCGGATAATTTGCCATTTCTCTCGGAATTAATTCCTATTTCCAAC 84996  
QY 41 TyrIysGlyPheIleArgMetThrGlyPheCysLysThrIysIle----- 55  
||||| ::  
Db 84997 TACAAGGGGTTCTGCGATAGGTGCTGTTCTGTAAAGCAAGGT-ATGTTGATCTCTT 85055  
QY 55 ----- 55  
Db 85056 TCATCTGTGTGATGATATGCTTCTGTGACACATGTCGCTTTGCTTAAAGGCT 85115  
QY 55 ----- 55  
Db 85116 TTAAGATGAGATTGAGTTTACTGCATCTATCTAGTTGCTTGGTGAATATATATAGACC 85175  
QY 55 ----- 55  
Db 85176 TTGGTGATGATGCTTATGACTAATGCTGTCAGATATCATTTTCTTGTGATTAATTC 85235  
QY 56 -----ProAlaSPHleMetAlaAlaLeuGluProIle 66  
||||| ::  
Db 85236 ACTCATGATCTTATTTATGATAGATACCGCTGAGCTCACTGCTGCTTAAGCCTAAT 85295  
QY 67 LysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeuGlyThrGluMetCysLys 86

```

Db 85296 AAGGTAATATGACGAGCTGTTAAAGCCTATGCAATTCACCTTGGCAACAAATGGCAA 85355
      |||||||
QY 87 LysileuAlaHisglyIleLysThrLeuHisLeuTyrThrLeuAsnMetIleLysSer 106
      |||||||
Db 85356 AAGATTTTGGCCATGGAATCACTCCCTCATCTCTACATGACATGAGCGTGGCAAAATCA 85415
      |||||||
QY 107 AlaLeuAlaIleLeuMet----- 112
      |||||||
Db 85416 GCTATTGGGATATTAAATGTTTCTCTCATTTCTCTTAATTTGGTTTCTGTGTAGTAA 85475
      |||||||
QY 113 -----As 113
      ||
Db 85476 TATTTAACACATGATTTGTATACCGGAGAACTAACAGTGTGCTCAATTTGCAGAA 85535
      |||||||
QY 113 nLeuGlyLeuIleGluLeuSerLysValSerArgSerLeuProTTPAArgProAlaAs 133
      |||||||
Db 85536 CTTGTGCTGATGTAGTCAAAATTTCTGTTCTCTACCTTGGAGAGCGCCGTGCAA 85595
      |||||||
QY 133 nValPheArgValLysGluAspValArgProIlePheTrp----- 146
      |||||||
Db 85596 TGTTCGCTACTAAGAGAGATGTGCCCCAATTTCTGTAACTGTCATTTTGTCTC 85655
      |||||||
QY 146 ----- 146
      ||
Db 85656 TTGCTCGAGATATTATACAGATCTTAACTTAAAGATATATTGGCGGAGAAATTAATCT 85715
      |||||||
QY 147 -----AlaAsnArgProLysSerTyrIleSerArgThrI 158
      |||||||
Db 85716 GTTACTTCATTTTATCTTTGACGAGCAAAACCGTCCAAAGAGCTACATATCTAGACAA 85775
      |||||||
QY 158 IegIlyTrpAspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyA 178
      |||||||
Db 85776 AAGGCTGAAGACTCTCCACATGAGAGCTGGGGTATTTACACAGTGCAGATACAGTA 85835
      |||||||
QY 178 IalLeuSerAspTyrGln----- 183
      |||||||
Db 85836 CACTTCGATTTATACAGTACGACACATATTATACAGTTTTCGACATATCGACTTAA 85895
      |||||||
QY 183 ----- 183
      ||
Db 85896 TATTTTGAACCTGAAATTCGAAAGGCTGTTTTCAGTTTGCCCTTCGAATTTGCTTAA 85955
      |||||||
QY 183 ----- 183
      ||
Db 85956 ATTGATTTGCATATTGTTATTAGCCCTACATATACCTGTAGAGAGCTGCTGCT 86015
      |||||||
QY 183 ----- 183
      ||
Db 86016 TGCTTGCATCTCTTGCCTGTTTTCATTCGCCAATGAGTAGTGTAAATGCAAGT 86075
      |||||||
QY 184 -----PheMetArgProArgIleAlaArgAspLysLysLeuValGluGluTrpAlaValP 201
      |||||||
Db 86076 GTTGAATGTTTGGCGCCCAAAAGAGCGTACAAAGAGCTTCAGCAAGAAATGGTGTGTC 86135
      |||||||
QY 201 rIleLysSerValGluAspIleTyrGluArgPheArgLeuTyr----- 215
      |||||||
Db 86136 CACTGAAAAGCATTAAGATGTCAAGAGCT---AAGCTTACTTTACGACACTCGTA 86191
      |||||||
QY 215 ----- 215
      ||
Db 86192 CAATTTGTTTGCATCTTTCCTCATACTACCTGCACTTTCATGTATAAAACTGTCA 86251
      |||||||
QY 216 -----CysLeuGlyLysLeuArgSerAsnProTTPSerGluT 228
      |||||||
Db 86252 CAATTCAGAAATTCAAAGAGCTCTGCATTTGGAACCTTAAAAAGCAGCCCATGTGTGAAT 86311
      |||||||
QY 228 euAspGlyLeuGlnProGluThrLysIleIleAsnGluLeuGluLysIleAsnThrL 248
      |||||||
Db 86312 TAGATGCACTCCAGCCAGAGCAAAAGATCAATTAATGTAGCAACTCGGAAAAATCAACTCA 86371
      |||||||
QY 248 yuGlyPheLeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProT 268
      |||||||

```

```

Db 86372 ACGGTTTCTGACCATTAATATGACCAACCATCATGCAATGACGCCAAATCGATTCTCAG 86431
      |||||||
QY 268 hVal----- 269
      |||
Db 86432 CATTTGGTAAGAAAATCAACTGATGATGACAAAGATACATTAACGTTCCAAAAAAA 86491
      |||||||
QY 269 ----- 269
      ||
Db 86492 AAATCTATTGACCTTGATTCATCTGTCACTTTCTTAAAAACCTATTAAACCTCTAG 86551
      |||||||
QY 269 ----- 269
      ||
Db 86552 AGCTTGTCTTATAGCTTATACATTAATGTTTAAATTGCAATGGGAATATCTAGCTT 86611
      |||||||
QY 270 -----GlyTrpGlyGluProGlyGlyTyr 277
      |||||||
Db 86612 ATTAACAAAATGTATTAATGCAATTTGATCTTTCAGAGATGGGGTGTCTGTGTTTAC 86671
      |||||||
QY 278 ValTyrGluLysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuVal 297
      |||||||
Db 86672 GTCTACAGAAAAGCTTATATGAGTGTCTGTCTTCAAAAGATTAAGTACACACTTGTG 86731
      |||||||
QY 298 AspLysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrp 317
      |||||||
Db 86732 GAGAAATCCAAAGCTTTTCTCTTATACCTACATGCGCGGTGAACAAATCAGAGATTGG 86791
      |||||||
QY 318 LysSerAsnValGlyGlnThrAspValAsnAlaValThrTrpGlyValPheProAlaLys 337
      |||||||
Db 86792 GTATCAAAACCGGTGAGCTCGATGTGATGATGATCACTTCTGGGGAGTGTCCACGTTAG 86851
      |||||||
QY 338 GluIleIleGlnProThrIleValAspProValSerPheAsnValTrpLysAspGluAla 357
      |||||||
Db 86852 GAGGTATTCAACCCACATGCTGATCCAGCCAGTTTCAAAAGTCTGGAAAGCAGAGG 86911
      |||||||
QY 358 PheGluIleTrpSerArgGlyTrpAlaSerLeuTyrProGluAspGluAlaSerArgLys 377
      |||||||
Db 86912 TTTGAGATTTGGTCAAGAGCTGGGCTTAACCTTGTACCCAGAGATACCCCTTGTAAAG 86971
      |||||||
QY 378 LeuValGluGluValGly----- 383
      |||||||
Db 86972 TTGCTGAGAGAGGTAA--GAATCTTCTCTTTTCTCTCTCTCTTTCATGACCAAGAT 87030
      |||||||
QY 384 -----GlySerHis 386
      |||
Db 87031 TGCTTAGCTTAAGATGATGATGATTCATTCGAATGCAATGACAGGTGAACACACTAC 87090
      |||||||
QY 387 PheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAspLeuPheAlaValPheAla 406
      |||||||
Db 87091 TATTGTGTAAGCTTATGTGACACAAATTAATCAATGATGATATCTCTGTCTTGTCT 87150
      |||||||

```

RESULT 11  
 AF441241 1785 bp mRNA linear SYN 01-AUG-2002  
 LOCUS AF441241  
 DEFINITION Synthetic construct chimeral mRNA, complete cds.  
 ACCESSION AF441241  
 VERSION AF441241.2 GI:22040812  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

2 (bases 1 to 1785)  
 Roje,S., Raymond,R.K., Applling,D.R. and Hanson,A.D.  
 Direct Submission  
 Submitted (29-Oct-2001) Horticultural Sciences, University of

REFERENCE Florida, Hull Road, Gainesville, FL 32611, USA  
3 (bases 1 to 1785)  
AUTHORS Roje,S., Raymond,R.K., Appling,D.R. and Hanson,A.D.  
TITLE Direct Submission  
JOURNAL Submitted (01-AUG-2002) Horticultural Sciences, University of  
Florida, Hull Road, Gainesville, FL 32611, USA  
REMARK Sequence update by subletter  
COMMENT On Aug 1, 2002 this sequence version replaced gi:19550442.  
FEATURES  
Location/Qualifiers  
source  
1..1785  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
1..1785  
/note="yeast-Arabidopsis hybrid methylenetetrahydrofolate  
reductase"  
/codon\_start=1  
/transl\_table=11  
/product="Chimeral"  
/protein\_id="AAL91367.2"  
/db\_xref="GI:22040813"  
/translation="MKITEKLEHQRHRTSGKPTYSFEYFVKTTQGVONLDRMDHMYE  
ASLPQIDITWAGGRLSHLSTDLVATAQSVGLGTCMLTCTNMPISMDALENA  
YHSGCONILALGDDPRDAENWTFEGGPOYAKDLIKRISKYGDHPAIGVAGECH  
PELPKDYKIDLELYKOKIDAGDPITTOCFYDNDNTKMSQVRAAGMDVPITIGIM  
PTTYAALFRRANQGISIPHFSSRIDPIKDDDELVRDGTGLTYEMGOKILDSGY  
SHDITMKLEKAPLMLTERKULIPTSERNAHPLAVLPKRLNKRKNEEVRFLW  
ANRPSYSIRTKGMDPFHGWSDSHSAAYSTUSDQFAPRKDRKLDQEEVPLKS  
IEDVQEKELCIGNLSSPWSLELDQPEYTKIINDELKINSNGFLITNSQPVAA  
KSDSPAIGWGGGGVYOKAYLEFCSKDKIDLVEKSKAPSIITMAVKNSENVSN  
TGESDVNAVTVGVPFAKEVIQPTIVDPAFVWMDAEFIIMSWMANLYPEDDPSRKL  
LEEVKSYLVAVLDNNYINGDIFSVA"  
BASE COUNT 506 a 415 c 424 g 440 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.92e-115 Length: 1785  
Score: 1411.00 Matches: 258  
Percent Similarity: 78.00% Conservative: 61  
Best Local Similarity: 63.08% Mismatches: 84  
Query Match: 64.52% Indels: 6  
DB: 12 Gaps: 3  
US-09-720-451-6 (1-408) x AF441241 (1-1785)  
Oy 4 TLevalthrGlnleupheryrAsprhrAspIleleuLyspheValAsnAspCyArg 23  
Db 556 ATCATCTACAGATGTTTACGATGTTGATATTCATCACTGCGTTCCCAAGTTAGA 615  
Oy 24 GlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAspTyrGly 43  
Db 616 GCTGCGGCGATCGACGTCGCAATTATCCCGGATCATGCCGATCAGTACCTACGCGGCC 675  
Oy 44 PheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMetAlaIleu 63  
Db 676 TTCTTGAGAAAGGCGCCAAATCTCCATCTCCATCAATTTCTCTCCCGCAT 735  
Oy 64 GluProIleLysAspAngIuAlaValLysAlaTyrGlyIleHisLeuGlyThrGlu 83  
Db 736 GATCCTTTCAGAGCAGTACGAGTGTGCTGATATCCGAACTGAACATTATGCTGGAA 795  
Oy 84 MetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyrThrLeuAsn 102  
Db 796 ATGTCGCAAAAATGCTCGACAGTGTGTTCTCTCACTGCAATCTACACCATGAMC 855  
Oy 103 MetCysLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIle---GluGluSerLys 121  
Db 856 TTGGAAGAAAGCGCTCTCATGATTCGGAAGATGACATTCACCTACGAGATTCAGAG 915  
Oy 122 ValSerArgSer-----LeuProTTPArgArgProAlaAsnValPheArgVal 137  
Db 916 TTCATATCACATCATTCGCGCGTGTGTCATGAGAAATCTTGGAATCCAAAGCGTAAA 975  
Oy 138 LysGluAspValAlaArgProIlePheTrpAlaAsnArgProLysSerTyrIleSerArgThr 157

Db 976 AACGAGAACTCGACCTTATCTTGGCGAAACCGTCCAAAGACTCATATCTAGAAC 1035  
Oy 158 TLeGlyTrpAspGlnProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGly 177  
Db 1036 AAGGCGTGAATGACTTCCCATGATGAGCGTGGGTGATTCACAGATGACATGACATGAC 1095  
Oy 178 AlaIleSerAspTyrLysIlePheMetArgProAlaArgAspLysLysLeuValGluGlu 197  
Db 1096 ACACCTTGGATTTATCAGTTTCGGCGCCAAAGAGCGTCCGAAAGAGCTTCACAGAA 1155  
Oy 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217  
Db 1156 TGGGTGCTCCCATGAAAGCATTTGAGATGTTCAAGAGAAATTCAAAGAGCGCTGCATT 1215  
Oy 218 GlyLysLeuArgSerAsnProTrpSerGluLeuAspGlyLeuGlnProGluThrLysIle 237  
Db 1216 GGAAGACTTAAAGAGCGCCATGGTCTGAAATTAGATGACTCCAGCCAGACAAAGATC 1275  
Oy 238 TLeAsnGluGlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnPro 257  
Db 1276 ATAAATGAGCACTCGAAGAAATCAACATCCACGAGTTCCTGACCATTAATAGCAACCA 1335  
Oy 258 AlaValAsnGlyLysSerAspSerProThrValGlyTyrGlyLysIleTyrGlyTyr 277  
Db 1336 TCACTCATACACCAAAATCCGATTCACAGCTATTGATGGGTGGTCTGCTGCTGCTAC 1395  
Oy 278 ValTyrGlnLysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuVal 297  
Db 1396 GTCTACACGAAACCTTATCTGAGTCTCTGTTCAAGGATTAAGTTAGACACACTGTG 1455  
Oy 298 AspLysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrp 317  
Db 1456 GAGAAATCCAAAGCTTTCTCTCTATACACTACATGCGCGTGAACAAATCAGAAATGC 1515  
Oy 318 LysSerAsnValGlyGlnThrAspValAsnAlaValThrTrpGlyValPheProAlaLys 337  
Db 1516 GTATCAAAACACCGTAGTCCGATGATGATGATGATGATGATGATGATGATGATGATG 1575  
Oy 338 GluIleIleGlnProThrIleValAspProValSerPheAsnValTyrLysAspGluAla 357  
Db 1576 GAGCTTATTCACACGACGATCGGATCCAGCCATTTCAAAAGCTCTGGAAGAGCAAGCG 1635  
Oy 358 PheGluIleTrpSerArgLysTrpAlaSerLeuTyrProGluAspGluAlaSerArgLys 377  
Db 1636 TTTGAGATTTGGTCAAGAGCTGGGCTGTAATCTGTACCCAGAAATGACCTTACAGAAAG 1695  
Oy 378 LeuValGluGluValGlyLysSerHisPheLeuValSerLeuValAspAsnAspTyrIle 397  
Db 1696 TTGCTCGAGGAGGTGAAGAAACAGCTACATTTGTGTAACCTTAGAGCAACATTAATC 1755  
Oy 398 AsnGlyAspLeuPheAlaValPheAla 406  
Db 1756 AATGTGATATCTCTGCTGCTTGTCT 1782  
RESULT 12  
AX069359 1971 bp DNA linear PAT 25-JAN-2001  
LOCUS AX069359  
DEFINITION Sequence 23 from Patent WO0102600.  
ACCESSION AX069359  
VERSION AX069359.1 GI:12579222  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1971)  
Yuan,C.S.  
TITLE Detection of analytes using attenuated enzymes  
AUTHORS Patient: WO 0102600-A 23 11-JAN-2001;  
JOURNAL GENERAL ATOMICS (US)  
FEATURES  
source  
1..1971  
Location/Qualifiers



Db 847 GAGGTGATGAGCAATCAAGACAAGATGCTGCCATCCGCAACTATGCGATGAGCTG 906  
 Qy 81 G1yThcGlUmeCysLysLysLleLeuAlaHisGly---LleLysThrLeuHisLeuTyr 99  
 Db 907 GCCGTGAGCTGTGGCAGAGAGCTTGTGGCACTGGCTGGCCAGAGGCTCCACTTCTAC 966  
 Qy 100 ThrLeuAsnMetGluLysSerAlaLeuAlaLleLeuMetAsnLeuGlyLleuLleGluGlu 119  
 Db 967 ACCCTCAACCGCGAGATGGCTACCCACAGAGTGTGAGAGCCCTGGGGATGTGACACTGAG 1026  
 Qy 120 SerLysValSerArgSerLeuProTrrArgArgProAlaAsnValPheArgValLysGlu 139  
 Db 1027 GACCC---AGGCGTCCCTTACCGCTGGGCTCACTGAGCCACCCCAAGCCGAGAGGAA 1083  
 Qy 140 AspValArgProLlePheThrAlaAsnArgProLysSerTyrLleSerArgThrLleGly 159  
 Db 1084 GATGTAGCTCCCATCTTCTGGGCTCCAGACCAAGAGATTACACTACCGTACCAGAGAG 1143  
 Qy 160 TrpAspLlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 179  
 Db 1144 TGGAGACAGTTCCTTAACGCGCGCTGGGCAATTCTTCCCTGGGAGAGCTG 1203  
 Qy 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197  
 Db 1204 AAGAGACTACCTCTTCACTGAGAGCAAGTCCCCCAAGAGAGAGAGCTGTAGATG 1263  
 Qy 198 TrpAlaValProLeuLysSerValGluAspLleTyrGluArgPheArgLeuTyrCysLeu 217  
 Db 1264 TGGGGGAGAGAGCTGACAGCAAGCAAGTCTTGAAGCTTGTCTTACCTCTCG 1323  
 Qy 218 G1yLysLeuArgSerAsn-----ProTrrSerGluLeuAspGly 230  
 Db 1324 GAGAGACCAAAACCGAATGTCACAAAGAGACTTCCCTGCCCTGGAGAGAT---GAGCC 1380  
 Qy 231 LeuGlnProGluThyLysLleLleAsnGluGluLeuGlyLysLleAsnThrLysGlyPhe 250  
 Db 1381 CTGGCGGCTGAGACAGCTGTGAGAGAGAGCTGCGGGTGAGCAACCGCAGGCGCATC 1440  
 Qy 251 LeuThrLleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrValGly 270  
 Db 1441 CTCACGTCACACTGACAGCCCAACATCAAGGAGAGCGCTCTCCAGCCCATGTGGGC 1500  
 Qy 271 TrpGlyGlyProGlyLysLysValTyrValTyrGlnLysAlaTyrValLysLysPheCysSerLys 290  
 Db 1501 TGGGGGCCAGCGGGGCTATGTCTCCAGAGGCGCTTGAAGTTTTCACCTCCCGC 1560  
 Qy 291 G1uLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309  
 Db 1561 GAGACAGCGAGACACTTCTGCAGAGTGTCAAGAGAGTACAGAGCTCCGGTTAATTACAC 1620  
 Qy 310 AlaValAsnLysAspLysSerTrrLysSerAsnValGlyGlnThrAspValAsnAlaVal 329  
 Db 1621 CTGTCTCAAGTGAAGGAGTAAACATCAACCAATGCCCTCCAGTCAAGCCGAAAGCTGTG 1680  
 Qy 330 ThrTrpGlyValPheProAlaLysGluLleLleGlnProThrLleValAspProValSer 349  
 Db 1681 ACTTGGGCGATCTTCCCTGGGAGAGATCATCCAGCCACCGTAGTGGATCCCGTACG 1740  
 Qy 350 PheAsnValTrpLysAspLysAlaPheGluLleTrrPserArgGlyTrpAlaSerLeuTyr 369  
 Db 1741 TTATGTCTTGGAGAGACAGGCGCTTGGCCCTGTGGATGAGCGGTGGGAAAGCTGTAT 1800  
 Qy 370 ProGluAspLysLysArgLysLeuValGluGluValGlyLysSerHisPheLeuVal 389  
 Db 1801 GAGAGAGAGTCCCGTCCCGACATCATCCAGTACATCCAGCAACTACTTCCGTGTC 1860  
 Qy 390 SerLeuValAspAsnAspTyr---LleAsnGlyAspLeuPheAlaValPheAlaAsp 407  
 Db 1861 AACCTGTGTGACATGACTTCCACTGGAACAATGCTGTGGCAGGTGTGGAGAGAC 1917  
 RESULT 14  
 AX050441  
 LOCUS AX050441 2187 bp DNA linear PART 12-JAN-2001

DEFINITION Sequence 1 from Patent WO0071754.  
 ACCESSION AX050441  
 VERSION AX050441.1 GI:12226648  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Johnson, W.G. and Stenroos, E.S.  
 TITLE Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors  
 JOURNAL University of Medicine and Dentistry of New Jersey (US)  
 FEATURES  
 source  
 1..2187  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 479 a 655 c 616 g 437 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5-5e-77 Length: 2187  
 Score: 979.00 Matches: 188  
 Percent Similarity: 63.728 Conservative: 79  
 Best Local Similarity: 44.878 Mismatches: 138  
 Query Match: 44.764 Indels: 14  
 Gaps: 7  
 US-09-720-451-6 (1-408) x AX050441 (1-2187)  
 Qy 1 AlaArgValLleValThrGlnLeuPheTyrAspTrpAspLlePheLeuLysPheValAsn 20  
 Db 667 GCCGATTTCATCAACAGCAGAGCTTCTTGAAGCTGACACATTTCTCCGTTGTGTAAG 726  
 Qy 21 AspCysArgGlnLleGlyLleThrCysProLleValProGlyLleMetProLleAsnAsn 40  
 Db 727 GCATGCACACCAATAGGAGCATCACTCCCATGCTCCCGGAGTCTTCCATCCAGAGGC 786  
 Qy 41 TyrlsGlyLleLleArgMetThrGlyPheCysLysThrLysLleProAlaAspLleMet 60  
 Db 787 TACCATCTCCCTGGCGAGCTGTGAAAGCTGTCCAAAGCTGGAGGTGACAGAGAGAG 846  
 Qy 61 AlaLleGluLysProLleLysAspAsnGluAlaValLysAlaTyrGlyLleLysLeu 80  
 Db 847 GACGTGATGAGCAATCAAGACAAGATGCTGCATCCGCACTATGAGATGAGAGCTG 906  
 Qy 81 G1yThcGlUmeCysLysLysLleLeuAlaHisGly---LleLysThrLeuHisLeuTyr 99  
 Db 907 GCCGTGAGCTGTGGCAGAGAGCTTGTGGCACTGGCTGGCCAGAGGCTCCACTTCTAC 966  
 Qy 100 ThrLeuAsnMetGluLysSerAlaLeuAlaLleLeuMetAsnLeuGlyLleuLleGluGlu 119  
 Db 967 ACCCTCAACCGCGAGATGGCTACCCACAGAGTGTGAGAGCCCTGGGGATGTGACACTGAG 1026  
 Qy 120 SerLysValSerArgSerLeuProTrrArgArgProAlaAsnValPheArgValLysGlu 139  
 Db 1027 GACCC---AGGCGTCCCTTACCGCTGGGCTCACTGAGCCACCCCAAGCCGAGAGGAA 1083  
 Qy 140 AspValArgProLlePheThrAlaAsnArgProLysSerTyrLleSerArgThrLleGly 159  
 Db 1084 GATGTAGCTCCCATCTTCTGGGCTCCAGACCAAGAGATTACACTACCGTACCAGAGAG 1143  
 Qy 160 TrpAspLlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 179  
 Db 1144 TGGAGACAGTTCCTTAACGCGCGCTGGGCAATTCTTCCCTGGGAGAGCTG 1203  
 Qy 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197  
 Db 1204 AAGAGACTACCTCTTCACTGAGAGCAAGTCCCCCAAGAGAGAGAGCTGTAGATG 1263  
 Qy 198 TrpAlaValProLeuLysSerValGluAspLleTyrGluArgPheArgLeuTyrCysLeu 217

[illegible]

JOURNAL	Submitted (17-MAY-1994) Rima Rozen, Pediatrics, Human Genetics and Biology, McGill University - Montreal Children's Hospital, 2300 Tupper St., Montreal, Quebec H3H 1P3, Canada
REFERENCE	4 (bases 1 to 2196)
AUTHORS	Leclerc,D., Sibani,S. and Rozen,R.
TITLE	Direct Submission
JOURNAL	Submitted (29-OCT-1999) Pediatrics, Human Genetics and Biology, McGill University - Montreal Children's Hospital, 2300 Tupper St., Montreal, Quebec H3H 1P3, Canada
REMARK COMMENT FEATURES	Amino acid sequence update by submitter On Nov 2, 1999 this sequence version replaced gi:945022.  Location/Qualifiers
SOURCE	1..2196 /organism="Synthetic construct" /db_xref="taxon:32630"
gene	1..2196 /gene="MTHR"
misc_feature	1..8 /gene="MTHR" /note="EcoRI linker" 13..1983 /gene="MTHR" /note="Method: conceptual translation with partial peptide sequencing"
CDS	/codon_start=1 /product="methyltetrahydrofolate reductase" /transl_table=1 /protein_id="AA04440.2" /db_xref="GI:6139053" /translation="MVNENARNSSLNPCLSGSSGSRKSDSRCSPTGDPDRPHHE LRMKRRRLSGLDKWSELEPPRPRTAGAVNLISRPDMAAGCPLIDVWYGGDGG SDKTSSTMIASTAVNTCGLETIIHMCCRRLEETIGHLHKAKQLGKNIMALRGR IGDWEEEGGFNAVVDLVHIRSEFDYEDICVAGYPKHPEKGSFEADIKHLKEV SAGDAFTITOLFEPADFFFRVKACTDMGITCPGIPIFGYHSLSROLVKLSLEV POELKDVIIEPKDDAAILRNQIGLAIVSLCOELASGLVGLHFYTILNREMAETEVLK RLGWATDEPRRLPDWALSHPHKKREEVRPIFMASRKSYTYRQEWDEPNRGWSNV SSPAFGLKLYLLFYTKSKSPKEELLMMGEELTSEAYPEVFLYLISEPNRGHKVY TCLPWNDEPLAETSLLKEELLRYNRGGITLTINQPINKGPSDDPIYGMPSGGYVF OKALEFTFSKETAEALLQVLRKYELARNVHLVANKENTTNAPELPNATWTGIFPG RETIQTPVDFVSFMFKDEAFALMIERMGVLVNEESPSRTIIOYIHDFVNLVNDN DEPDNCIMQVEVDTELELNRPQNARETEAP"
variation	1298 /gene="MTHR" /note="E429A; reported by Weisberg et al (1998) Mol. Genet. Metabol. 64, 169-172" /replacement="a" 2196 /gene="MTHR" /note="18 A nucleotides" /evidence-experimental
polya_site	BASE COUNT      482 a                  618 g                  439 t ORIGIN
Alignment Scores:	
Pred. No.:	5.53e-77                  Length:                  2196
Score:	979.00                  Matches:                 188
Percent Similarity:	63.72%                  Conservative:            79
Best Local Similarity:	44.87%                  Mismatches:            138
Query Match:	44.76%                  Indels:                 14
DB:	12                  Gaps:                    7
US-09-720-451-6 (1-408) x HSU09806 (1-2196)	
OY	1 AlaargvalillevallhrcgllleupheyrasptlrAspillepheleulyspheValasn 20        :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::   Db 676 GCCGATTTCATCATCCGCACGCTTTTCTTGAGGCCGACACATCTTCCGCTTTGTGAAG 735     ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::   OY 21 AspCysArgGlnIlleQlyllerThcCysProileValProGlylleMetProileasn 40     ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::   Db 736 GGATGACCGACGATGGGCACTTGCCCCCANCGTCCCAGGATCTTCCATCCAGGGC 795     ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::   OY 41 TyrLyseglyPhelIeargethriglyPheCysLysThrLyalleProAlaIspleket 60     ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

Db 796 TACCACCTCCCTTCGGACGCTGTGAAGCTGTCCAAAGCTGGAGGTCACAGAGATCAAG 855  
QY 61 AlaAlaLeuGluProIleuIleuValAspAsnGluGluAlaValIleuValIleuIleuLeu 80  
Db 856 GACGTGATGTGACCAATCAAGACACGATGCTGCACATCCGCAATCTGTGCATCGAGCTG 915  
QY 81 GlyThrGluMetCysIleuValIleuAlaHisGly---IleuValIleuHisIleuValIleu 99  
Db 916 GCCGTGAGCCTGTGCTGCAGAGGCTCTGCGCAGTGGCTTGGTCCAGGCTCCACTTCTAC 975  
QY 100 ThrLeuAsnMetGluIleuValIleuValIleuValIleuValIleuValIleuValIleu 119  
Db 976 ACCCTCAACCGCGAGATGCTACACAGAGTGTGTAAGCGCTGGGATGTGACTGAG 1035  
QY 120 SerIleuValSerArgSerLeuProTyrArgArgProAlaAsnValIleuValIleuValIleu 139  
Db 1036 GACCCC---AGCGTCCCTTACCTGGGCTCTCAGTCCCAAGCCCAAGCGCGAGAGAA 1092  
QY 140 AspValArgProIleuPheTyrAlaAsnArgProIleuValIleuValIleuValIleuValIleu 159  
Db 1093 GATGTACGTCCCATCTTCTGGGCTCCAGACCAAGATTCATCTACCTACCCAGAG 1152  
QY 160 TyrAspGluIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 179  
Db 1153 TGGGACGAGTTCCTTAACGCGCGCTGGGCAATCTCTCCCTGGCTTGGGAGCTG 1212  
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspIleuValIleuValIleuValIleu 197  
Db 1213 AAGGACTACTACTCTCTTCTGACGTAAGAGCAAGTCCCGCAAGAGAGAGCTGTGAAGATG 1272  
QY 198 TyrAlaValProLeuIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 217  
Db 1273 TGGGGGAGGAGCTGACAGCTGAAGCAAGTCTTGAAGTCTTGTGTTACTCTCTCG 1332  
QY 218 GlyIleuValArgSerAsn-----ProThrSerGluLeuAspGly 230  
Db 1333 GGAGAACCAACCGGAATGCTCAAAAGTACTTCCCTGCGCTGGAACGAT--GAGCCC 1389  
QY 231 LeuGluProGluThrIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 250  
Db 1390 CTGGCGGCTGAGACAGCCTGTGTAAGAGGAGCTGTGGGGTGAACCGCCAGGCGATC 1449  
QY 251 LeuThrIleuAsnSerGlnProAlaValAsnGlyGluIleuValIleuValIleuValIleuValIleu 270  
Db 1450 CTCACCATCTACCTACAGCCCAACATCAACGAGGAGCCGCTCCGACCCCATCGTGGCG 1509  
QY 271 TyrGlyGlyProGlyIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 290  
Db 1510 TGGGGCCCCAGCGGGGCTATGCTCTCCAGAAAGCCTACTTAGATTCTTCACTTCCCGC 1569  
QY 291 GluIleuValAspAlaLeuValAspIleuValIleuValIleuValIleuValIleuValIleuValIleu 309  
Db 1570 GAGACAGCGGAGCACTTCTCAAGTGTGAAGTACAGAGCTCCGGTTAATTACAC 1629  
QY 310 AlaValAsnIleuValAspGlySerTyrIleuValIleuValIleuValIleuValIleuValIleuValIleu 329  
Db 1630 CTGTCAATGTGAAGGTGTAACATCACCAATGCCCTGAACCTGCACCGCAATGCTGTC 1689  
QY 330 ThrTyrGlyValPheProAlaValIleuValIleuValIleuValIleuValIleuValIleuValIleu 349  
Db 1690 ACTTGGGGCATCTTCCCTGGCGAGAGATCATCCAGCCACCGTACGTCCCGTCAAC 1749  
QY 350 PheAsnValIleuValAspGluAlaPheGluIleuValIleuValIleuValIleuValIleuValIleuValIleu 369  
Db 1750 TTCATGTTCGAAAGAGCAGCGCTTGGCTGTGATGAGCGGTGGGAAAGCTGTAT 1809  
QY 370 ProGluAspGluValSerArgIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 389  
Db 1810 GAGGAGAGAGCGCGCTCCGACCATCATCCAGTACATCCACGACACAACTACTCTGTC 1869  
QY 390 SerIleuValAspAsnAspTyr---IleuAsnGlyAspLeuPheAlaValIleuValIleuValIleuValIleu 407  
Db 1870 AACCTGTGAGACATCTTCCCATCTGACAACTGCTCTGGCAGGTGTGGAAGAC 1926

Search completed: February 15, 2003, 00:16:56  
Job time : 3670 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 12, 2003, 17:03:09 ; Search time 58 Seconds

(without alignments)  
2157.312 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187

Sequence: 1 ARVITOLFDTDFLEKFN.....VSLVDNDYINGDLPAVADF 408

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xml  
-MODE=frame+ p2n.model -DEV-xml  
-O=/cgn2\_1/USFTO.spool/US09720451/runat\_05022003\_074809\_2384/app\_query.fasta\_1.583  
-Db=Issued\_Patents\_NA -OFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US09720451.cgn1.1.17 @runat\_05022003\_074809\_2384 -NCPD=6 -ICPD=3  
-NO\_XLPPY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database:

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	979	44.8	1971	4	US-09-347-878-23
2	979	44.8	2187	4	US-09-318-448-1
3	979	44.8	2219	3	US-08-738-000-3
4	979	44.8	2219	4	US-09-258-928-3
5	979	44.8	2220	3	US-08-738-000-1
6	979	44.8	2320	4	US-09-258-928-1
7	148.5	6.8	1449	4	US-08-858-207A-54
8	148.5	6.8	28882	4	US-08-961-527-140
9	138	6.3	4848	4	US-08-951-957A-1
10	138	6.3	4848	4	US-08-951-957A-4
11	138	6.3	4848	4	US-08-951-957A-6
12	109.5	5.0	1556	2	US-08-687-355A-3

13	109.5	5.0	1556	4	US-09-407-367-3	Sequence 3, Appli
14	104.5	4.8	14311	4	US-08-646-695-1	Sequence 1, Appli
15	104.5	4.8	14311	4	US-08-646-695-7	Sequence 7, Appli
16	104.5	4.8	14311	5	PCT-US96-06053-1	Sequence 1, Appli
17	104.5	4.8	14311	5	PCT-US96-06053-7	Sequence 7, Appli
18	100	4.5	38584	4	US-09-453-702B-50	Sequence 50, Appli
19	97.5	4.5	8220	2	US-08-568-459A-11	Sequence 11, Appli
20	97.5	4.5	8220	2	US-08-487-826B-11	Sequence 11, Appli
21	97.5	4.5	8220	4	US-09-210-288-11	Sequence 11, Appli
22	97.5	4.5	19124	2	US-08-487-826B-13	Sequence 13, Appli
23	93.5	4.3	1900	4	US-08-965-762-22	Sequence 22, Appli
24	93.5	4.3	1900	4	US-09-911-927-22	Sequence 22, Appli
25	93.5	4.3	1900	4	US-09-911-927-22	Sequence 24, Appli
26	93.5	4.3	1900	4	US-09-911-882-22	Sequence 24, Appli
27	93.5	4.3	1900	4	US-09-911-882-24	Sequence 24, Appli
28	93	4.3	2800	4	US-09-056-285A-7	Sequence 24, Appli
29	92.5	4.2	1654	4	US-09-221-017B-326	Sequence 326, App
30	89	4.1	1366	1	US-08-670-354-5	Sequence 5, Appli
31	89	4.1	1366	4	US-09-320-424-5	Sequence 5, Appli
32	89	4.1	1366	5	PCT-US96-10895-5	Sequence 5, Appli
33	89	4.1	2144	4	US-08-876-798A-1	Sequence 1, Appli
34	89	4.1	2144	4	US-08-876-798A-3	Sequence 3, Appli
35	87.5	4.0	1320	1	US-08-290-448A-75	Sequence 75, Appli
36	87.5	4.0	1320	1	US-08-290-448A-75	Sequence 75, Appli
37	87.5	4.0	1320	1	US-08-175-669A-75	Sequence 75, Appli
38	87.5	4.0	1320	4	US-08-461-939B-75	Sequence 75, Appli
39	87.5	4.0	1320	4	US-08-461-939B-75	Sequence 75, Appli
40	87.5	4.0	1320	4	US-09-393-529-1	Sequence 1, Appli
41	87.5	4.0	2835	1	US-08-750-532-2	Sequence 2, Appli
42	87.5	4.0	4765	1	US-08-750-532-8	Sequence 8, Appli
43	87.5	4.0	4765	4	US-08-894-818B-7	Sequence 7, Appli
44	87.5	4.0	4765	4	US-09-445-472-5	Sequence 5, Appli
45	87	4.0	3619	4	US-08-377-503-1	Sequence 1, Appli

#### ALIGNMENTS

RESULT 1  
US-09-347-878-23  
Sequence 23, Application US/09347878C  
Patent No. 6376210  
GENERAL INFORMATION:  
APPLICANT: Yuen, Chong  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES  
FILE REFERENCE: 25885-1651  
CURRENT APPLICATION NUMBER: US/09/347,878C  
CURRENT FILING DATE: 1999-07-06  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 23  
LENGTH: 1971  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Human methyltetrahydrofolate reductase (MTHFR)  
PUBLICATION INFORMATION: gene: exons 1-8  
DATABASE ACCESSION NUMBER: AF105977/Genbank 1-11  
US-09-347-878-23

Alignment Scores:  
Pred. No.: 7.42e-110  
Score: 979.00  
Length: 1971  
Best Similarity: 63.728  
Matches: 188  
Best Local Similarity: 44.87%  
Mismatch: 138  
Query Match: 44.76%  
Indels: 14  
DB: 4  
Gaps: 7

US-09-720-451-6 (1-408) x US-09-347-878-23 (1-1971)

QY 1 AAlaryVallleValhrgInleuPhetYrAsprhAspIlePhetleuYsPhavaIaAn 20  
Db 664 GCGATTTCATCACTACGCGCGCTTTCTTGTAGCGTCAACATTCCTCCGCTTGTGAAG 723

Qy	21	AspCygaArgInIleGlyIleThrCysProIleValProGlyIleMetProIleAsn	40
Db	724	GCATGACACCGACATGGGATCATCTGGCCCATGTCGCCGGATCTCCCATCCAGGCC	783
Qy	41	TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProIleAspIleMet	60
Db	784	TACCACCTCCCTTCGGCACCTGTGAACTGTCCAGCTGGAGGTGCACAGGATCAAG	843
Qy	61	AlaIleLeuGlnProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu	80
Db	844	GAGCATTTGACGCAATCAAGACACAGATGCGTCATCCGCACTATGGCATCGAGCTG	903
Qy	81	GlyThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyr	99
Db	904	GCCGTAGGCTGTGGCAAGGACCTTGTGGCAGAGTGTGGTGCCAGGCGTCACCTCTAC	963
Qy	100	ThrLeuAsnMetGluLysSerAspIleLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGln	119
Db	964	ACCCCTCAACCGGACGATGGCTACACACAGAGGTGCTGAAGCGCTGGGATGTGACTGAG	1022
Qy	120	SerLysValSerArgSerLeuProTyrArgArgProIleAlaAsnValPheArgValLysGln	139
Db	1024	GACCC---AGCGTCGCCCTACCTGGGCTCTCAGTGGCCCAACCGCCAGAGGGA	1081
Qy	140	AspValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGly	159
Db	1081	GATGTACGTCCTCATTTCTTGCGCTCCACGACCAAGATCATCTCCGTACCCAGAG	1140
Qy	160	TyrAspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeu	179
Db	1141	TGGGACGATTTCCCTTAACGCGCGGTGGGCAATCTCTTCCCGCTTTGGGAGAGTG	1200
Qy	180	SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGluGln	197
Db	1201	AAGATATTACTACTCTTCTTACTCTAAGAGCAATGCCCAAGAGAGAGCTGCTGAAGATG	1267
Qy	198	TyrAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu	217
Db	1261	TGGGGGAGAGACTACACCGACAGCAAGTGTCTTGAAAGTCTTGTTCTTACCTCTCG	1320
Qy	218	GlyLysLeuArgSerAsp-----ProTyrSerGluLeuAspGly	230
Db	1321	GGAGAACCAACCGGAATGGTCAACAAGTACTGTGCTGCGCTGCATGAGCAT--GAGCCC	1377
Qy	231	LeuGlnProGluThrLysIleIleAsnGluGlnLeuGlnLysIleAsnThrLysGlyPhe	250
Db	1378	CTGGCGGCTGAACACCGCTCTCTGAAGAGAGACTGTCTGGGTGAACCCCGAGGCATC	1437
Qy	251	LeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspSerProThrValGly	270
Db	1438	CTCAACATCACTCAACGCCCAATCAACGGGAGCGGTCTCCGACCCCACTGGGGG	1497
Qy	271	TyrGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys	290
Db	1498	TGGGGCCCAAGGGGGGTATGCTTCCAGAAAGCCCTACTTAGAGTTTCTCACTCCGC	1555
Qy	291	GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet	309
Db	1558	GAAACACGGGAAGCACTTCTCAAGTCTGAAGAAAGTACGACGCTCCGGTAAATATCCAC	1617
Qy	310	AlaValAsnLysAspGlySerTyrLysSerAsnValGlyGlnThrAspValAsnAlaVal	329
Db	1618	CTTGTCATAGTGAAAGGGAAGAAACATCAACCAATGCCCTCGAAGTCAGCGGATGTCTGC	1677
Qy	330	ThrTyrGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer	349
Db	1678	ACTTGGGGCATTTCTCCCTGGGGCGAGATCATCTACGCCCCAGTAGAGATCCCGTAC	1737
Qy	350	PheAsnValTyrLysAspGluAlaPheGluIleTyrSerArgGlyTyrAlaSerLeuTyr	369
Db	1738	TTATATCTTGTCGAAGAGACGAGCGCTTGGCTGTGATTTGAGCGGTGGGAAGCTGTAT	1797

[illegible]



QY 218 GLYLeuArgSerAsn-----ProtpserGluLeuAspGly 230  
|||:  
Db 1333 GGAGAACCAACCGGAAATGCTACAAAGTGCCTGCTCCCTGGAAACGAT---GAGCCC 1389  
QY 231 LeuGlnProGluThrLysIleIleAsnGluIleGluLysIleAsnThrLysGlyPhe 250  
|||:  
Db 1390 CTGGCGCTGAGACCAAGCTGCTGTAAGAGAGAGCTGCTGGGGTGAACCCCGCAGGCGATC 1449  
QY 251 LeuThrIleAsnSerGlnProAlaValAsnGluLysSerAspSerProThrValGly 270  
|||:  
Db 1450 CTCACCATTAACCTACAGCCCAACATCAACGGGAAGCCGCTCTCCGACCCCATCGTGGGC 1509  
QY 271 TrpGlyGlyProGlyGlyTyValTyGlnLysAlaIleValGluPhePheCysSerLys 290  
|||:  
Db 1510 TGGGGCCCGCAGCGGGGCTATGTCTTCCAGAAAGCCCTAGCTTGAAGCTTTTCCACTCCGC 1569  
QY 291 GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyMet 309  
|||:  
Db 1570 GAGACAGCGGAAAGCACTTCTGCAAGTGTGAAGAAAGTACAGCTCCCGGTATATTACAC 1629  
QY 310 AlaValAsnLysAspGlySerThrLysSerAsnValGlyIleThrAspValAsnAlaVal 329  
|||:  
Db 1630 CTGTCAATGTGAAGGAGGAAACATCAACATGCCCCCTGAATGACGCGGAATGCTGTC 1689  
QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349  
|||:  
Db 1690 ACTGGGGCATCTTCCCTGGGGGAGAGATCATCCAGCCACCGTAACTGATCCCGTCAGC 1749  
QY 350 PheAsnValThrLysAspGluAlaPheGluIleThrSerArgLysThrAlaSerLeuTy 369  
|||:  
Db 1750 TTCATCTTGTGAAGAGAGAGGCTTTTCCCTGCTGATGTAAGCGGCTGGGAAAGCTGAT 1809  
QY 370 ProGluAspGluAlaSerArgLysLeuValGluIleValGlySerHisPheLeuVal 389  
|||:  
Db 1810 GAGAGAGATCCCGCTCCCGCCGACCATCATCCAGTACATCCAGCAGCATCTTCCCTGCTG 1869  
QY 390 SerLeuValAspAsnAspTy---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407  
|||:  
Db 1870 AACCTGTGTGACAAATGACTTCCACTGGAACAACTGCTTGGCAGGTGTGGAAAGAC 1926

RESULT 4  
US-09-258-928-3  
; Sequence 3, Application US/09258928  
; Patent No. 6218120  
; GENERAL INFORMATION:  
; APPLICANT: KOZEN, Rima  
; APPLICANT: GOYETTE, Philippe  
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE  
; FILE REFERENCE: 04844/005002  
; CURRENT APPLICATION NUMBER: US/09/258, 928  
; CURRENT FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 08/738, 000  
; PRIOR FILING DATE: 1997-02-12  
; PRIOR APPLICATION NUMBER: GB 9410620. 0  
; PRIOR FILING DATE: 1994-05-26  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2219  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (13)...(1983)  
US-09-258-928-3

Alignment Scores:  
Pred. No.: 9,01e-110 Length: 2219  
Score: 979.00 Matches: 188  
Percent Similarity: 63.72% Conservative: 79  
Best Local Similarity: 44.87% Mismatches: 138

Query Match: 44.76% Indels: 14  
DB: 4 Gaps: 7  
US-09-720-451-6 (1-408) x US-09-258-928-3 (1-2219)  
QY 1 AlaArgValIleValThrGlnLeuPheThrAspThrIlePheLeuLysPheValAsn 20  
|||:  
Db 676 GCCGATTTCATCAACAGCAGCTTTTCTTGGAGCTGACATCTTCCCTTGTGAG 735  
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn 40  
|||:  
Db 736 GCATGCCACCACTATGGGCATCACTTCCCATCTGCCATCCGCACTATGGATGAGCTG 795  
QY 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrIleProAlaAspIleMet 60  
|||:  
Db 796 TACCACTCCCTTGGCAGCTTGTGAAGCTGTGCCAAGCTGAGCCAGCAAGAGTAAG 855  
QY 61 AlaAlaLeuGluProIleLysAspAsnGluIleAlaValLysAlaTyGlyIleAlaLeu 80  
|||:  
Db 856 GAGCTGATTAGCCATCAAGAACAAGATGCTGCCATCCGCACTATGGATGAGCTG 915  
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTy 99  
|||:  
Db 916 GCCGTGACCTGTGCGCAGAGCTGTGGCAGTGTGCTGTCAGAGCCCTCCACTTAC 975  
QY 100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyIleGluGlu 119  
|||:  
Db 976 ACCCTCAACCGCAGATGGCTACACAGAGTGTGAAGCGCCTGGGGATGTGACTGAG 1035  
QY 120 SerLysValSerArgSerLeuProThrArgArgProAlaAsnValPheArgValLysGlu 139  
|||:  
Db 1036 GACCCC---AGGGTCCCTTCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1092  
QY 140 AspValArgProIlePheThrAlaAsnArgProLysSerTyIleSerArgThrIleGly 159  
|||:  
Db 1093 GATGTAGCTCCCATCTTCTTGGGCTTCCAGACCAAGAAATTTACATCTACCGTACCCAGAG 1152  
QY 160 TrpAspGlnThrProHisGlyArgTrpGlyAspSerCysAsnProSerTyGlyAlaLeu 179  
|||:  
Db 1153 TGGGACGAGTTCCCTAAGCCGCGCTGGGGCAATTCCTTCCCTGGGAGACTG 1212  
QY 180 SerAspTyGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu 197  
|||:  
Db 1213 AAGCACTACTACTCTTCTTACCTGGAAGACAAAGTCCCAAGAGAGAGCTCTAAGATG 1272  
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyGlnArgPheArgLeuTyCysLeu 217  
|||:  
Db 1273 TGGGGGAGAGAGCTGACAGTGAAGCAAGTGTCTTGAAGTCTTGTACCTCTCG 1332  
QY 218 GlyLysLeuArgSerAsn-----ProtpserGluLeuAspGly 230  
|||:  
Db 1333 GGAGAACCAACCGGAAATGCTACAAAGTGCCTGCTCCCTGGAAACGAT---GAGCCC 1389  
QY 231 LeuGlnProGluThrLysIleIleAsnGluIleGluLysIleAsnThrLysGlyPhe 250  
|||:  
Db 1390 CTGGCGCTGAGACCAAGCTGCTGTAAGAGAGAGCTGCTGGGGTGAACCCCGCAGGCGATC 1449  
QY 251 LeuThrIleAsnSerGlnProAlaValAsnGluLysSerAspSerProThrValGly 270  
|||:  
Db 1450 CTCACCATTAACCTACAGCCCAACATCAACGGGAAGCCGCTCTCCGACCCCATCGTGGGC 1509  
QY 271 TrpGlyGlyProGlyGlyTyValTyGlnLysAlaIleValGluPhePheCysSerLys 290  
|||:  
Db 1510 TGGGGCCCGCAGCGGGGCTATGTCTTCCAGAAAGCCCTAGCTTGAAGCTTTTCCACTCCGC 1569  
QY 291 GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyMet 309  
|||:  
Db 1570 GAGACAGCGGAAAGCACTTCTGCAAGTGTGAAGAAAGTACAGCTCCCGGTATATTACAC 1629  
QY 310 AlaValAsnLysAspGlySerThrLysSerAsnValGlyIleThrAspValAsnAlaVal 329  
|||:  
Db 1630 CTGTCAATGTGAAGGAGGAAACATCAACATGCCCCCTGAATGACGCGGAATGCTGTC 1689



QY 370 progilysagluinalasercqylsleuvalglugllylgsrthspheulval 389  
 |||||  
 |||||  
 Db 1810 GAGGAGAGATCCCGCTCCGACCATCATCTCCAGTACATCCAGCAACTCTCTCGGTC 1865  
 |||||  
 |||||  
 QY 390 SerleuValaspasnspTY---lleasnglaspleuPhealavalPhealasp 407  
 |||||  
 |||||  
 Db 1870 AACCTGTGTGACAAAGATCTCCACATGGAGCAACTCCCTCTGGCAGGTGGGGAAGC 1926  
 |||||  
 |||||

RESULT 6  
US-09-25

Sequence 1, Application US/09258928  
Patent No. 6218120  
GENERAL INFORMATION:  
APPLICANT: BOZEV, Rima  
APPLICANT: GORETTE, Philippe  
TITLE OF INVENTION: GNA FOR HUMAN METHYLENETETRAHYDROFOLATE  
TITLE OF INVENTION: REDUCTASE  
FILE REFERENCE: 04844/005002  
CURRENT APPLICATION NUMBER: US/09/258,928  
PRIORITY FILING DATE: 1999-03-01  
PRIOR APPLICATION NUMBER: 08/778,000  
PRIORITY FILING DATE: 1997-02-12  
PRIOR APPLICATION NUMBER: GB 9410620.0  
PRIORITY FILING DATE: 1994-05-26  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 2220  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1980)  
IS-09-258-928-1

### Alignment Scores:

Pred. No.:	9,02e-110	length:	2220
Score:	979.00	Matches:	188
Percent Similarity:	63.72%	Conservative:	79
Best local Similarity:	44.87%	Mismatches:	138
Query Match:	44.76%	Indels:	14
DB:	4	Gaps:	7

US-09-720-451-6 (1-408) x US-09-258-928-1 (1-2220)

QY 1 AlaATGValIleValThrGlnIlePheItyrAspIhrAspIlePheLeuIysPheValAsn 20  
Db 676 GCCGATTCATCATACGACGACCTTTCTTTGAGGCTGACACATCTTCGCGTTTGGAAG 735  
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40  
Db 736 GCATCGACCGACAAATGGGATCACTTCCCGCATCGTCCCGGGATCTTCCCATCCAGGCG 795  
QY 41 TyrIysGlyPheIleArgMetThrGlyPheCysIsthThrIleProIleAspIleMet 60  
Db 796 TACCACCTCCCTTCGGCAGACTTGTAAGCTGTCCAGAGCTGGAGTGGCCACAGAGCAAG 855  
QY 61 AlaAlaIleGlnProIleIysAspAsnGlnGluAlaValIleAlaTyrGlyIleHisIleu 80  
Db 856 GACGTGATGACCATCAAGAGACAGATGCTGCATCCGACATCCCACTATGGCATCGAGCTG 915  
QY 81 GlyThrIleMetCysIysIleIleValIleAsnIleGly--IleIsthThrIleuHisIleuTyr 99  
Db 916 GCGCTAGGCGCTGTGCAGAGAGCTTCTGGCAGCTGGTGGTCCAGAGCTCCACCTCTCAC 975  
QY 100 ThrLeuAsnMetGlnIysSerAlaIleAlaIleIleMetAsnIleuGlyIleGlnGlu 119  
Db 976 ACCCTCAACCGGAGATGGCAACACAGAGGTGTGAAGGCCCTGGGGAGTGGACTGAG 1035  
QY 120 SerIysValSerArgSerLeuProTPrpArgIleProIleAsnValPheArgValIysGlu 139  
Db 1036 GACGCC--AGCGTTCCTCTTACCTCTGGCTCTCAGAGGCCACCCAGAGCGCCAGAGGA 1092

QY	140	AspValArgProIlePheTrpAlaAsnArgProLysSerTrpIleSerArgThrIleGly	159
Db	1093	GAGTATGCGCCCATCTTCTGGCGCTCCAGACCAAGAAATTCATCTACGCTACCAGAG	1153
QY	160	TrpAspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu	179
Db	1153	TGGGAGAGTATCCCAACAGCGCGCTGGGCAATCTCTTCCCTCGCTTTGGGAGCTG	1212
QY	180	SerAspTyrGln-----PheMetArgProAlaArgAspLysIleLeuValGluGlu	197
Db	1213	AAGAGCTACTACTCTTCTTACTCTGAAGACAGTCCCCCAAGGAGGAGCTGCTGAAGATG	1272
QY	198	TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu	217
Db	1273	TGGGGGAGAGACTGTACAGCAAGAGCAAGTGTCTTGAAGCTCTTGTCTTTCACCTCTG	1332
QY	218	GlyIleLeuArgSerAsn-----ProTyrSerGluLeuAspGly	230
Db	1333	GGAAGAACCAACCGAATGGTCACCAAGTACTTGTGCTGCCCTGTGAACAT---GAGCCC	1389
QY	231	LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe	250
Db	1390	CTGGCGCGCTGAAGACCAGCTGTCTGAAGAGAGCTGTGCGGTGAACCCCGAGGACATC	1449
QY	251	LeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspSerProThrValGly	270
Db	1450	CTACCATCACTACAGCCCAACATCAACGGGAAACCGCTCTCCGACCCATCCATCGGGC	1509
QY	271	TrpGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluIlePheCysSerLys	290
Db	1510	TGGGCGCCCGAGCGGGGGGTATGCTTCCAGAAAGGCTACTTAGAGTTTTCACCTTCCGC	1569
QY	291	GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet	309
Db	1570	GAGACAGCGGAGACACTTCTCTCAAGTGTCTGAAGAGTACAGACCTCCGGTATTATCCAC	1629
QY	310	AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal	329
Db	1630	CTTGTCAAGTGAAAGGAGGAAACATCACCAATGCGCTGAGTAACTGACCGCAATGTGTC	1689
QY	330	ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer	349
Db	1690	ACTTTGGGGCATCTTCCGTGGGCGAGAGATCATCCAGCCACCGTATGGATCCCGTCAGC	1749
QY	350	PheAsnValTrpLysAspGluAlaPheGluIleTyrPheSerArgGlyTrpAlaSerLeuTyr	369
Db	1750	TTCATGTCTTGGAAGAGAGAGCGCTTTCGCCCTGTGATGATGACCGGTGGGAAAGCTGTAT	1809
QY	370	ProGluAspGluAlaSerArgLysLeuValGluGluValGlyGlySerHisAspLeuVal	389
Db	1810	GAGAGAGAGTCCCGCTCCCGCACCATCATCCAGTACATCCACGACACACTCTCTCGTGC	1869
QY	390	SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp	407
Db	1870	AACCTGTGGAAATGACTTCTCCACTGTGAACATCGCTCTTGTGGCAGGTGTGCAAGAC	1926

## RESULT 7

US-06-036-20/A-54  
 ? Sequence 54, Application US/08858207A  
 ? Patent No. 6348328  
 ?  
 ? GENERAL INFORMATION:  
 ?  
 ? APPLICANT: Black, Michael  
 ?  
 ? APPLICANT: Hodgson, John  
 ?  
 ? APPLICANT: Knowles, David  
 ?  
 ? APPLICANT: Nicholas, Richard  
 ?  
 ? APPLICANT: Stodola, Robert  
 ?  
 ? TITLE OF INVENTION: No. 6348328e1 Compound  
 ?  
 ? NUMBER OF SEQUENCES: 552  
 ?  
 ? CORRESPONDENCE ADDRESS:  
 ?  
 ? ADDRESSEE: SmithKline Beecham Corporation  
 ?  
 ? STREET: 709 Swedeland Road  
 ?  
 ? CITY: King of Prussia  
 ?  
 ? STATE: PA  
 ?



Db 15502 AATGCTGATACAGCA 15488

RESULT 9

US-08-955-957A-1

Sequence 1, Application US/08955957A

Patent No. 6312920

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Rostock Jr., Paul R.

TITLE OF INVENTION: SAM Operon

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/955,957A

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: P-10162

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4848 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 986..2209

NAME/KEY: CDS

LOCATION: 2241..3341

US-08-955-957A-1

Alignment Scores:

Pred. No.: 4.68e-06 Length: 4848

Score: 138.00 Matches: 30

Percent Similarity: 43.64% Conservative: 18

Best Local Similarity: 27.27% Mismatches: 62

Query Match: 6.31% Indels: 0

DB: 4 Gaps: 0

US-09-720-451-6 (1-408) x US-08-955-957A-1 (1-4848)

OY 5 ValThrGlnLeuPheThyAspThrAspIlePheLeuLysPheValAsnAspCysArgGln 24

DB 3872 ATCAGCGAGATGTTCTTCGCGGTGAGGACTATCTCGCGGTGCGGACCGGGTCCGCCGCC 3931

OY 25 IleGlyIleThrCysProIleValProGlyIleMetProIleAsnAspTyrLysGlyPhe 44

DB 3932 GCCGCGCTGTCGACCCGCGGTATCCCGGATCATGCCGCCACCGAGCTCCGCGCATC 3931

OY 45 IleArgMetThrGlyPheCysLysThrLysIleProIleAspIleMetAlaIleLeuGln 64

DB 3992 GCGCGCTTCCGCGGACTTTCACGCGCCACTTCCCGGAGAGGTCTCGCGCGCGCGTGGAG 4051

OY 65 ProIleLysAspAsnGlnGluAlaValLysAlaTyrGlyIleHisLeuGlyThrGluMet 84

Db 4052 GCCGCCCGCGCACCCGCCGAGGAGACCCGATCGGGGTGATGATACGCCACCGCATG 4111

OY 85 CysLysIleIleValAlaHisGlyIleLysThrIleHisLeuTyrThrLeuAsnMetGlu 104

DB 4112 GCCGCGCGCTCTCTGCGCGGAGAGGCCGCCCGGAGCTGCACTCATCATCACCCTCC 4171

OY 105 LysSerAlaLeuAlaIleLeuMetAsnLeu 114

DB 4172 ACCGCGACGTTGAGATCATCCACCGGAACATC 4201

RESULT 10

US-08-955-957A-4

Sequence 4, Application US/08955957A

Patent No. 6312920

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Rostock Jr., Paul R.

TITLE OF INVENTION: SAM Operon

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/955,957A

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: P-10162

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 4848 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 3338..4255

US-08-955-957A-4

Alignment Scores:

Pred. No.: 4.68e-06 Length: 4848

Score: 138.00 Matches: 30

Percent Similarity: 43.64% Conservative: 18

Best Local Similarity: 27.27% Mismatches: 62

Query Match: 6.31% Indels: 0

DB: 4 Gaps: 0

US-09-720-451-6 (1-408) x US-08-955-957A-4 (1-4848)

OY 5 ValThrGlnLeuPheThyAspThrAspIlePheLeuLysPheValAsnAspCysArgGln 24

DB 3872 ATCAGCGAGATGTTCTTCGCGGTGAGGACTATCTCGCGGTGCGGACCGGGTCCGCCGCC 3931

OY 25 IleGlyIleThrCysProIleValProGlyIleMetProIleAsnAspTyrLysGlyPhe 44

DB 3932 GCCGCGCTGTCGACCCGCGGTATCCCGGATCATGCCGCCACCGAGCTCCGCGCATC 3931

```

Oy      45  ILeIrgvethrGlyPheCysLysThrIlySIIleProAlaSPrlMetAlaIalaIeuIu 64
          |||
Db      3992  GCGGCGCTTCCGCCAGCTTTTCCACGCCACCTTCCCGGAAGTCTCGCGCGGCGCTGGAG 4051
          |||

Oy      65  ProlIeYsapsngInguAlaValIySAlaTyrgIlyIleHISleuG1YThrgIuMet 84
          ::| |||
Db      4052  GCGCGCGCGGCAACCGCGCGGAGGAGACACCGCATGCGGGGTGAGTACGCCACCGCCATG 4111
          ::| |||

Oy      85  CysLysIlySIIleIleAlaHISgIlyIleLysThrIleuHISleuTyThrIleuAsnMetIu 104
          ::| |||
Db      4112  GCGCGCGCGGTCTCGCGCGGAGGCGCGCGCGGACCTGACATCATCATCTCAACCGCTCC 4171
          ::| |||

Oy      105  LysSerAlaIleAlaIleuMetAsnIeu 114
Db      4172  ACCGCGACGTTCGAGATCCACCGGAACATC 4201

RESULT 11
US-08-955-957A-6
Sequence 6, Application US/08955957A
Patent No. 6312920

GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Rosteck Jr., Paul R.
TITLE OF INVENTION: SAM Operon
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,957A
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: P-10162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 4848 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-955-957A-6

Alignment Scores:
Pred. No.: 4.68e-06 Length: 4848
Score: 138.00 Matches: 30
Percent Similarity: 43.64% Mismatches: 18
Best Local Similarity: 27.27% Mismatches: 62
Query Match: 6.31% Indels: 0
DB: 4 Gaps: 0

US-09-720-451-6 (1-408) x US-08-955-957A-6 (1-4848)

Oy      5  ValThrgInIeuPheIyrrAspThIrrAspIIlePheIeuLysPheValAsnAspCysArgIn 24
          ::| |||
Db      3872  AUCAGCAGAGUUCUUCGCGGUCGAGACUACUUCGCGGCGGAGACCGCGGUCGCGGCG 3931
          ::| |||

```

```

QY      25   IlegIleThrCysProIleValIProglyIleMeProIleAsnAspTyrLysGlyphe 44
           |||
Db       3932 GCCGCUCGUCACCCCGGGACUCCCCGGAGAUCAUCCGCCACCGAGUCGGCAGAUC 3991
           |||
QY      45   IleArgMetThrGlyPheCysLysTyrLysIleProAlaASPILeMetAlaLeuGlu 64
           |||
Db       3992 GCGGCGUUCGCCGAGCUUUUCCACGGCACCUUCCCGGAAGUGUCUGCGGGCGGUGSAG 4051
           |||
QY      65   ProIleLysAspAsngIugIuaIaValLysAlaTyrglyIleHisLeuglyThrglunet 84
           ::::
Db       4052 GCCGCCGCCGACACCGCGCCGAGGAGACACCGCAUCCGGGUGAGUACGCCACGCCAUG 4111
           ::::
QY      85   CysLysLysIleLeuAlaHSLgIyIleLysThrLeuHISLeuTyThrLeuasnMetLu 104
           ::::
Db       4112 GCCGCCGCGCUCGCCCGCGGAAGCGCCCGCGGAGUCCGACUAUCAUCCUACCGCUC 4171
           ::::
QY      105  LysSerAlaLeuAlaIleLeuMetAsnLeu 114
           ::::
Db       4172 ACCGCGAGCUUGGAGAUCCACCGGAAACAUC 4201
           ::::

RESULT 12
US-08-687-355A-3
; Sequence 3, Application US/08687355A
; Patent No. 5989834
; GENERAL INFORMATION:
; APPLICANT: Synaptic Pharmaceutical Corporation
; TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE
; NUMBER OF INVENTION: Y/PEPTIDE YY (YY) RECEPTORS AND USES THEREOF
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,355A
; FILING DATE: No. 5989834ember 26, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPM/MAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1556 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOCHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 211..1353
; US-08-687-355A-3

Alignment Scores:
Pred. No.:          0.00227         Length:        1556
Score:              109.50          Matches:        81
Percent Similarity: 35.42%          Conservative: 49
Best Local Similarity: 22.07%       Mismatches:    136
Query Match:        5.01%           Indels:        101
Df:                  2               Gaps:          20

```



```

QY 128 Trp-----ArgArgProAlaAsnValPheArgValLysGlu----- 139
    |||
Db 727 TGGGGTGTACAGCGCTGCTGCTGGCAAGTCCCTTGCCATCTTCGGGAGTACTACTATT 786
QY 140 AspValArgProIlePhe-----TyrPalaAsnArgProLys 151
    |||
Db 787 GAGATTATTCTCGACTTGTGACATTGTAGCCCTGTACTGAGAAATGGCCGGGAGAGAGAG 846
QY 152 Ser-----TyrIleSerArgThr-----Ile 158
    |||
Db 847 AGTGTGTACAGGTACAGTGTACAGCCTTTCACCTGTCTACTGTTTTCCTGTG 906
QY 159 GlyTyrPArgIleTyrProHisGlyArg---TyrGlyAspSerCysAsn---ProSerTyr 176
    |||
Db 907 GGCATCATATCTTCTCTCCATACCCGGATCTGGAGTAAAGCTAAAGAACCCAGTTAGTCT 966
QY 177 GlyAlaLeuSerAspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGlu 196
    |||
Db 967 GGAAGTGCAGAGTACCATTTACATCAGCGAAGGACAAAGACCAAAATGCTGCTG--- 1023
QY 197 GluTyrPalaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCys 216
    |||
Db 1024 -----TGC 1026
QY 217 LeuGlyLysLeuArgSerAsnProTyrPheSerGluLeuAspGlyLeuGlnProGluThrLys 236
    |||
Db 1027 GTGGTAGTGGTGTGTGGCTGCTGCTGCCCTTCCTCAATGCTTCAATTCGTGTGAC 1086
QY 237 IleIleAsnGluGlnLeuGlnLysIleAsnThrLysGlyPheLeuThrIleAsnSerGln 256
    |||
Db 1087 ATCGACAGCAGCATGTCTGACCTGAGAGGATCAAAACATCATTCACCGTGTCCACAT 1146
QY 257 ProAlaValAsnGlyLysSerAspSerProThrValGlyTyrGlyGlyProGlyGly 276
    |||
Db 1147 ATTCGATGTGTCCACCTTCCTCCCAACCCCTTCTATGCTGTGATGAACAGCAAC--- 1203
QY 277 TyrValTyrGlnLysAlaTyrValGluPhePheCysSerLysGlyLysLeuAspAlaLeu 296
    |||
Db 1204 -----TACGAAAGACCTTCTCTCAGCCTTCGCTGAGCAGACAGGTGATGCAT 1257
QY 297 ValAspLysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySer 316
    |||
Db 1258 -----CACTCGAGAGGTGCCATGACCTTCAAGGCT---AAAAAGAACCTGGA 1302
QY 317 TrpLysSerAsnValGlyGlnThrAsp-----ValAsnAla----- 328
    |||
Db 1303 GTCAAAAGAACACATGGCTCCTGCTTTTTCAGAGGCCACCAACGCTGAAGATGC 1362
QY 329 -----ValThrTrp-----GlyValPhePro 335
    |||
Db 1363 TGTGAAGATACGTGGTAATTGCCGACAGAGTGGCCAACTGGTTAGGGAAGGTT--- 1419
QY 336 AlaLysGluIleIleGlnProThrIleValAspProValSerPheAsnValIleTrpLysAsp 355
    |||
Db 1420 TCTGGCTAGTGCATGCCACCTCCATGTGATGACCCCTAAAGACCAAGATGGAGGCC 1479
QY 356 GluAlaPheGluIleTyrPheSerArgGlyTyrPalaSerLeuTyrProGluAspGluAlaSer 375
    |||
Db 1480 CACGCGTATTGTTCTCTGGAAAAGTGGCTGGAGAAAT-----GAGGAGAAAATAAAC 1530
QY 376 ArgLysLeuValGluGluVal 382
    |||
Db 1531 AGATTGCTGTGGCGCAACGTT 1551

```

```

RESULT 14
US-08-646-695-1
; Sequence 1, Application US/08646695
; Patent No. 6186943
; GENERAL INFORMATION:
; APPLICANT: Rose, John K.
; TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
; NUMBER OF SEQUENCES: 44

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,695
FILING DATE: On Even Date Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14311 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 760..2025
FEATURE:
NAME/KEY: CDS
LOCATION: 2092..2886
FEATURE:
NAME/KEY: CDS
LOCATION: 2946..3632
FEATURE:
NAME/KEY: CDS
LOCATION: 3774..5306
FEATURE:
NAME/KEY: CDS
LOCATION: 5429..11755
US-08-646-695-1
Alignment Scores:
Pred. No.: 0.356 Length: 14311
Score: 104.50 Matches: 71
Percent Similarity: 34.06% Conservative: 54
Best Local Similarity: 19.35% Mismatches: 117
Query Match: 4.78% Indels: 125
Gaps: 17
US-09-720-451-6 (1-408) x US-08-646-695-1 (1-14311)
QY 80 LeuGlyThrGlnGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyr 99
    |||
Db 691 ATGGGACGAGACAAACAAACATATTATCATTTAAAGGCTCAGGAGAACTTAAACA 750
QY 100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlu 119
    |||
Db 751 GTAATCAAAATGTCTGTACAGTCAAGAAATCATTTGACAAACAGTCAATGATGTTCCAAA 810
QY 120 SerLysValSerArgSerLeuProTyrPargArgProAlaAsnValPheArgValLysGlu 139
    |||
Db 811 CTTCCTGCAAAATGAGGAT---CCAGTGAATATACCGGACGATTACTTAGAAAATCAAG 867
QY 140 AspValArgProIlePheThrPalaAsnArgProLysSer-----Tyr 153
    |||
Db 868 GAGATT---CCTCTTTAC---ATCAATACTACAAAAGTGGTGTGATGATGAAGATAT 921

```

```

QY 154 ILeSerArgThrIleGlyTyrPaspGlnTyrProHisGlyArgTyrGlyAspSerCysAsn 173
DB 922 GTTACCAAGGCGCTC-----AAATCCCGGAAT 948
QY 174 ProSer-----TyrGlyAlaLeuSerAspTyrGlnPheMet 185
DB 949 GTATCATCATACATGTCAACAGCTACTTGTATGGAGCATTAAGACATCCGG----- 1002
QY 186 ArgProAlaArgAspIysLysLeuValGluGluTyrPala----- 199
DB 1003 -----GTTAAGTTGATTAAGATTGGTCAAGTTTGGATAAATACATC 1044
QY 200 -----ValProLeuLysSerValGluAsp 207
DB 1045 GGGAAAGCAGGGGATACATCGGAATATTGGACCTTGTATCTTGAAGCCCTGGAGCGC 1104
QY 208 IleTyrGlu----- 210
DB 1105 GTACTTCCAGATGAGTATCGATGCTTCCAGAACCCAGCATGACAAATGGTTGGCT 1164
QY 211 -----ArgPheArgLeuTyrCysIleuGlyLysLeuArgSerAsnProTyrSerGlu--- 227
DB 1165 TTGTATCTACTTGGCTTATACAGAGGGGACAGACAAATGCTGATACAGAAAAAG 1224
QY 228 ---LeuAspGlyLeuGlnProGluTyrIleIleAsnGluGlnLeuGluIleAsn 246
DB 1225 CTCATGATGGCTGACAAATGATCAAAATGATCATGAAGCTTGAACCTCTGTG 1264
QY 247 ThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSer 266
DB 1285 CCGAAGGT-----CGTGACATT 1302
QY 267 ProThrValGlyTyrGlyProGlyLysValTyrGlnLysAlaTyrValGluPhe 286
DB 1303 TTGTAGTGT---TGGGGAATGACAGTAAATTAACAAAAATTTGCGCTGACGTGGACATG 1359
QY 287 Phe-----CysSerLysGluLysLeuAspAlaLeuValAsp 298
DB 1360 TTCTTCCACATGTCMAAAACATGATGCTGCTGATGACGAGACATATTGTTCC 1419
QY 299 LysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTyrLys 318
DB 1420 ACATTCMAAGATTTGCTGCGCATTTGCCAATTT-----GGACACCTCTGC 1464
QY 319 SerAsnValGlyGlnThrAspValAsnAlaValThrTyr----- 331
DB 1465 AAAATACCGGAATGTCTACAGAAAGATGTACAGACCTGATCTTGAACCGAAGTTGCA 1524
QY 332 -----GlyValPheProAlaLysGluIle-----IleGln 341
DB 1525 GATGAATGTCCTCAAAATGATCTTCCAGGCCCAAGAAATGACAAAGCCGATTCATCATG 1584
QY 342 ProThrIleValAsp-----ProValSerPheAsnValTyrLysAspGlu 356
DB 1585 CCTTATTGATGCGACTTTGGATTGCTTCTTAAGTCTCATATCTTCCGCAAAAACCT 1644
QY 357 AlaPheGluIleTyrSerArg 363
DB 1645 GCCTTCCACTCTGGGGGCA 1665

```

## RESULT 15

```

US-08-646-695-7/C
; Sequence 7, Application US/08646695
; Patent No. 6168943
; GENERAL INFORMATION:
; APPLICANT: Rose, John K.
; TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas

```

```

CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646 695
FILING DATE: On Even Date herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 14311 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-646-695-7

```

## Alignment Scores:

```

Pred. No.: 0.356 Length: 14311
Score: 104.50 Matches: 71
Percent Similarity: 34.06% Conservative: 54
Best Local Similarity: 19.35% Mismatches: 117
Query Match: 4.78% Indels: 125
DB: 4 Gaps: 17

```

US-09-720-451-6 (1-408) x US-08-646-695-7 (1-14311)

```

QY 80 LeuGlyThrGluMetCysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyr 99
DB 13621 ATAGGAGGAGAGCAACAAACCAATTTATTCATTAAAGCGTCAGAGAACTTTAAACA 13562
QY 100 ThrLeuAsnMetGluLysSerIleAlaLeuAlaIleLeuMetAsnLeuGlyIleGlu 119
DB 13561 GTAAATCAAAATGCTGTTCACAGTCAAGATCATTTGACAAACAGTCATATTGTTCCAAA 13502
QY 120 SerLysValSerArgSerLeuProTyrArgArgProAlaAsnValPheArgValLysGlu 139
DB 13501 CTTCCTGCAAAATGAGAT---CCAGTGAATATCCCGGACGATTAATTCAGAAATCAAG 13445
QY 140 AspValArgProIlePheThrPalaAsnArgProLysSer-----Tyr 153
DB 13444 GAGATT---CCTCTTAC---ATCAATATCTACAAAAAGTTGTCAGATCTAAGAGCATAT 13391
QY 154 ILeSerArgThrIleGlyTyrPaspGlnTyrProHisGlyArgTyrGlyAspSerCysAsn 173
DB 13390 GTTACCAAGGCGCTC-----AAATCCCGGAAT 13364
QY 174 ProSer-----TyrGlyAlaLeuSerAspTyrGlnPheMet 185
DB 13363 GTATCATCATACATGTCAACAGCTACTTGTATGGAGCATTAAGACATCCGG----- 13310
QY 186 ArgProAlaArgAspIysLysLeuValGluGluTyrPala----- 199
DB 13309 -----GTTAAGTTGATTAAGATTGGTCAAGTTTGGATAAATACATC 13268
QY 200 -----ValProLeuLysSerValGluAsp 207
DB 13267 GGGAAAGCAGGGGATACATCGGAATATTGGACCTTGTATCTTGAAGCCCTGGAGCGC 13208
QY 208 IleTyrGlu----- 210

```

```
Db 13207 G T A C T T C C A G A T G A G T T A C G G A T C C T T C C A G A A C C A C G C A G A T G A C A A A T G C T T G C C T 13148
      ::::
QY 211 -----ArgPheArgLeuTyrCysLeuGlyLysLeuArgSerAsnProTrpSerGlu--- 227
      ||||| ::::| ::::| ::::| ::::|
Db 13147 T T G T A T C T A C T T G G C T T T A C A G A T G G G C A G A C A A A T G C C T G A A T T A C A G A A A A A A G 13088
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 228 ---LeuAspGlyLeuGlnProGluThrLysIleIleAsnGlnLeuGlnLeuGlnLysIleAsn 246
      ::::| ::::| ::::| ::::| ::::| ::::|
Db 13087 C T C A T G A T G G G C T G A C A A A T C A A T G C A A A A A T G A T G A C A G T T G A A C C T C T T G T G 13028
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 247 ThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsnGlyLulysSerAspSer 266
      ::::| ::::| ::::| ::::| ::::| ::::|
Db 13027 C C A G A A G G T ----- 13010
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 267 ProThrValGlyTyrPglYlyProGlyLylYrValIYrGlnLysAlaTyrValGluPhe 286
      ||| ||||| ||| ||| ::::|
Db 13009 T T T G A T G T G ---T G G G A A A T G A C A G T A T T A C A C A A A A A T T G T C G T G C A G T G G A C A T G 12953
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 287 Phe-----CysSerLysGluLysLeuAspAlaLeuValAsp 298
      ||| ::::| ::::| ::::| ::::| ::::|
Db 12952 T T C T T C C A C A T G T C A A A A A A C A T G A A T G T G C C T G T C A G A T A C G A C T A T T G T T C C 12893
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 299 LysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrpLys 318
      ::::| ::::| ::::| ::::| ::::| ::::|
Db 12892 A G A T T C A A A G A T T G C T G C A T T G G C A C A C A T T -----G G A C A C C T C T G C 12848
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 319 SerAsnValGlyGlnThrAspValAsnAlaValThrTrp----- 331
      ||| ::::| ::::| ::::| ::::| ::::|
Db 12847 A A A A T A A C C G A A T G C T A C A G A A G A T G T A A C G A C C T G A C T T G A A C C A G A A G T T G C A 12788
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 332 -----GlyAlaPheProAlaLysGluIle-----IleGln 341
      ::::| ::::| ::::| ::::| ::::| ::::|
Db 12787 G A T G A A A T G G T C C A A A T G A T G C T T C C A G G C C A A G A A A T T G A C A G A G C G A T C A T A C A T G 12728
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 342 ProThrIleValAsp-----ProValSerPheAsnValTrpLysAspGlu 356
      ||| ::::| ::::| ::::| ::::| ::::|
Db 12727 C C T T A T T G A T G A C T T T G G A T T G T C T C T A A G T C C A T A T T C T T C G C T A A A A A A C C C T 12668
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 357 AlaPheGluIleTrpSerArg 363
      ||||| ||| ::::|
Db 12667 G C C T T C C A C T T C T G G G G C C A A 12647
      ::::| ::::| ::::| ::::| ::::|
```

Search completed: February 12, 2003, 17:06:31  
Job time : 112 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 15:12:06 ; Search time 24 Seconds  
(without alignments)  
500.190 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187  
Sequence: 1 ARVIVQLEFDIDFLKFN.....VSLVNDYINGDLFAVFAADF 408

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfillseq1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	979	44.8	656	3	US-08-738-000-4
2	979	44.8	656	4	US-09-258-928-4
3	979	44.8	656	4	US-09-347-878-24
4	979	44.8	660	3	US-08-738-000-2
5	979	44.8	660	4	US-09-258-928-2
6	148.5	6.8	288	4	US-08-858-207A-324
7	138	6.3	305	4	US-08-955-957A-5
8	97.5	4.5	2710	2	US-08-568-459A-12
9	97.5	4.5	2710	2	US-08-487-826B-12
10	97.5	4.5	2710	2	US-09-210-288-12
11	97.5	4.5	3060	2	US-08-487-826B-14
12	97	4.4	422	4	US-08-646-695-2
13	97	4.4	422	5	PCT-US96-06053-2
14	91	4.2	347	4	US-08-965-762-23
15	91	4.2	347	4	US-09-911-927-23
16	91	4.2	347	4	US-09-911-882-23
17	88.5	4.0	530	4	US-08-943-714-12
18	87.5	4.0	397	1	US-08-290-448A-76
19	87.5	4.0	397	1	US-08-290-448A-76
20	87.5	4.0	397	1	US-08-175-069A-76
21	87.5	4.0	397	1	US-08-461-939B-76
22	87.5	4.0	397	4	US-08-464-000-76
23	87.5	4.0	616	4	US-09-393-529-2
24	86	3.9	381	2	US-08-687-355A-4
25	86	3.9	381	2	US-08-687-355A-6
26	86	3.9	381	4	US-09-407-367-4
27	86	3.9	381	4	US-09-407-367-6

28	85	3.9	574	4	US-09-134-001C-3141	Sequence 3141, Ap
29	85	3.9	1023	2	US-08-198-446B-4	Sequence 4, Appl
30	85	3.9	1023	2	US-08-870-693-4	Sequence 4, Appl
31	85	3.9	2368	2	US-08-198-446B-15	Sequence 15, Appl
32	85	3.9	2368	2	US-08-870-693-15	Sequence 15, Appl
33	84.5	3.9	653	2	US-08-713-928B-9	Sequence 9, Appl
34	83.5	3.8	348	1	US-08-169-950-2	Sequence 2, Appl
35	83.5	3.8	616	4	US-09-134-001C-4618	Sequence 4618, Ap
36	83.5	3.8	653	4	US-09-439-923-2	Sequence 2, Appl
37	83	3.8	593	1	US-08-296-362-2	Sequence 2, Appl
38	82.5	3.8	1285	1	US-07-582-945-2	Sequence 2, Appl
39	82.5	3.8	1285	2	US-08-453-141-2	Sequence 2, Appl
40	82.5	3.8	1285	2	US-08-293-314-2	Sequence 2, Appl
41	82	3.7	291	1	US-08-670-354-6	Sequence 6, Appl
42	82	3.7	291	4	US-09-320-424-6	Sequence 6, Appl
43	82	3.7	291	5	PCT-US96-10895-6	Sequence 6, Appl
44	82	3.7	918	2	US-08-825-558-6	Sequence 6, Appl
45	82	3.7	918	4	US-09-312-611-6	Sequence 6, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-738-000-4
Sequence 4, Application US/08738000
Patent No. 6074821
GENERAL INFORMATION:
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
TITLE OF INVENTION: REDUCTASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: KLAUBER & JACKSON
STREET: Continental Plaza - 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/738,000
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CA95/00314
FILING DATE: 25-MAY-1995
APPLICATION NUMBER: GB 9410620.0
FILING DATE: 26-MAY-1994
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 656 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-738-000-4
Query Match 44.8%; Score 979; DB 3; Length 656;
Best Local Similarity 44.9%; Pred.No. 5,1e-91;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;
```

QY 1 ARVIVQLEFDIDFLKFN.....VSLVNDYINGDLFAVFAADF 408

DB 222 ADFTITLFEADTFRRFKACIDMGTCTPVGIFPIQTHSLRQVKSKLEVPQIR 281

QY 61 AALEPIKNEEAAKAGIHIGTEMCKKIIANG-ITLHLTYLNNKSAIILNMLGIER 119

DB 282 DVIEPIKDNMAIRNIGIELAVSLQCELLASGLVGLHFTYLNREMAFTFVLKRLGWTE 341



```
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/738,000
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CA95/00314
FILING DATE: 25-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9410620.0
FILING DATE: 26-MAY-1994
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-738-000-2

Query Match          44.8%; Score 979; DB 3; Length 660;
Best Local Similarity 44.9%; Pred. No. 5,1e-91;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;

QY 1 ARVYVQLFPTDTIFLKFVNDRCQIGITCPVPIGMINNYKGFIMTGFCKKIPADIM 60
D 226 ADFTITQLFEADTFEFVYACQDMGTCTPVPDIPFGIHSRLQVLKSLKLEVPQETK 285
QY 61 AALEPIKNDNEAVKAYGHIHGTGEMCKLIAGH-ITLHLVTLNMEKSAIILNGLIEE 119
D 286 DVIEPIKNDNAIRNYGIELAVSLCQELLASGLVPGLHFTLNRMAATTEVLKRLGMWTE 345
QY 120 SKVSRSLPWRPANYRVKEDVRPIFWANRPKSYISRTIGMDQYPHGRWGDSCNPSYCAL 179
D 346 DP-RRLPFWALSAHPKRRREDVRPIFWASRPKSYIYRTQEMDEFPNGRWGSSSPAGEL 404
QY 180 SDYQ--FMRPRAROKKLYEEMAVPLKSVEDIYERFRLYCGKLRSN-----PMSELDG 230
D 405 KDYLFYFKSPKKEELKMGEEELTSEASVEFYLYLISGEPRNRNGHKVTCPLPMD-EP 463
QY 231 LQPETKIINEOLEKINTKGFLTINSQPAVNGEKSDPTVGMGGPGGYVYOKAYVEEFCCK 290
D 464 LAETSILKEELLRNNGIITINSQPNINGKPSDPIVGMGPGGYVFOKAYLEFPTSR 523
QY 291 EKLDALVDKCD-RTSLTYMAVNDGSKSVNGQTDVNAVTVGWGFPAKEIIQPTIYDPS 349
D 524 ETAEALLQVLKKYELRVNHLVNVKGENITNAPELQPNNAVTVGWGFPGREIIQPTIYDPS 583
QY 350 FNVWKEDEAFELIRSGMASLYPEDEASRKLYEEVGSHPFLVSLVNDY-INGDLPFAVEAD 407
D 584 FFWKDEAFALMIERKGLIEEESPKRIIYIHNTFLVNLVNDPFLDNCIMQVVED 642

RESULT 5
US-09-258-928-2
Sequence 2, Application US/09258928
Patent No. 6218120
GENERAL INFORMATION:
APPLICANT: ROZEN, Rima
APPLICANT: GOYETTE, Philippe
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
TITLE OF INVENTION: REDUCTASE
FILE REFERENCE: 04844/005002
CURRENT APPLICATION NUMBER: US/09/258,928
CURRENT FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 08/738,000
PRIOR FILING DATE: 1997-02-12
```

```
PRIOR APPLICATION NUMBER: GB 9410620.0
PRIOR FILING DATE: 1994-05-26
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 660
TYPE: PRT
ORGANISM: Homo sapiens
US-09-258-928-2

Query Match          44.8%; Score 979; DB 4; Length 660;
Best Local Similarity 44.9%; Pred. No. 5,1e-91;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;

QY 1 ARVYVQLFPTDTIFLKFVNDRCQIGITCPVPIGMINNYKGFIRMTGFCCKKIPADIM 60
D 226 ADFTITQLFEADTFEFVYACQDMGTCTPVPDIPFGIHSRLQVLKSLKLEVPQETK 285
QY 61 AALEPIKNDNEAVKAYGHIHGTGEMCKLIAGH-ITLHLVTLNMEKSAIILNGLIEE 119
D 286 DVIEPIKNDNAIRNYGIELAVSLCQELLASGLVPGLHFTLNRMAATTEVLKRLGMWTE 345
QY 120 SKVSRSLPWRPANYRVKEDVRPIFWANRPKSYISRTIGMDQYPHGRWGDSCNPSYCAL 179
D 346 DP-RRLPFWALSAHPKRRREDVRPIFWASRPKSYIYRTQEMDEFPNGRWGSSSPAGEL 404
QY 180 SDYQ--FMRPRAROKKLYEEMAVPLKSVEDIYERFRLYCGKLRSN-----PMSELDG 230
D 405 KDYLFYFKSPKKEELKMGEEELTSEASVEFYLYLISGEPRNRNGHKVTCPLPMD-EP 463
QY 231 LQPETKIINEOLEKINTKGFLTINSQPAVNGEKSDPTVGMGGPGGYVYOKAYVEEFCCK 290
D 464 LAETSILKEELLRNNGIITINSQPNINGKPSDPIVGMGPGGYVFOKAYLEFPTSR 523
QY 291 EKLDALVDKCD-RTSLTYMAVNDGSKSVNGQTDVNAVTVGWGFPAKEIIQPTIYDPS 349
D 524 ETAEALLQVLKKYELRVNHLVNVKGENITNAPELQPNNAVTVGWGFPGREIIQPTIYDPS 583
QY 350 FNVWKEDEAFELIRSGMASLYPEDEASRKLYEEVGSHPFLVSLVNDY-INGDLPFAVEAD 407
D 584 FFWKDEAFALMIERKGLIEEESPKRIIYIHNTFLVNLVNDPFLDNCIMQVVED 642

RESULT 6
US-08-858-207A-324
Sequence 324, Application US/08858207A
Patent No. 6348328
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: No. 6348328e1 Compounds
NUMBER OF SEQUENCES: 352
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858,207A
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017670
```



Db 274 TYFHATCNGERTKGYCRNDQVPTVY---FDYVPOYLK-----WFEEMADEFCRRKN 323  
QY 203 KSEVDIYERFR-----LYC-----LGKLR-----SNPW----- 225  
Db 324 KIKIDYKRCRGKDKEDKDRYCSRNGYDEKTKRAIGKRGKOCISCLYACNPFYDWIN 383  
QY 226 SELDGLQPETKLINEOLEKINTKGFLLTINSOPAVNGEKSD---SPTVWGCGPGGYVQKA 282  
Db 384 NKEQFQKOKKKYDEIKKYE-----NGASGSGRQKRDAGGTTTNYDG-----YEKK 431  
QY 283 Y-----VEFCSEKELDALVDCKDRT--SLITYAVNKGDSKMSNNGQDVNAV 329  
Db 432 FYDELINKSEYRTVDFLEKLSNEICTKYKDEEGGTFDKNVSD---STSGASGTNYE 487  
QY 330 TWGVFPAKEIIOP-----TIYDVSEFNWKMDEAFETWISRGWASLY---PEDEASRKLY 379  
Db 488 SGTGYFRSKYQCPGCPYCGVKKVNNGSSNMEWERNKNGCKSG--KLYEPKPREGTTITI 545  
QY 380 EEYGGSHFLVSLVDN---DYINGD 400  
Db 546 LKSGKHDIIEKLNKFCDEKNGD 569

RESULT 9  
US-08-487-826B-12  
; Sequence 12, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,826B  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned  
; REGISTRATION NUMBER: 29,655  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2710 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium falciparum  
US-08-487-826B-12

Query Match 4.5%, Score 97.5; DB 2; Length 2710;

Best Local Similarity 19.8%, Pred. No. 3.6;  
Matches 88; Conservative 59; Mismatches 142; Indels 155; Gaps 27;  
QY 56 PADIMAAI-EPIKNDENAVKAYGHLGTEM---CKLIANGIKTLIHLTYLNMKRSALALI 111  
Db 182 PSQICTMLARSPADIGDLYRGDLYLGNDPEIKORQOLENNLKI 226  
QY 112 MNLGIEESKVSRSI.PWRRPANVFRKEDVRFIFW-ANRPKSYISRTIGWDQYPHGRMGD 170  
Db 227 --FGKTYEKLNGAEARYGNDPEFFKLRED---WMTANR-----ETV-WKALTICNMGN 273  
QY 171 S-----CN-----PSYGLSDY--QFMPPARKKUIVEWAVPL----- 202  
Db 274 TYFHATCNGERTKGYCRNDQVPTVY---FDYVPOYLK-----WFEEMADEFCRRKN 323  
QY 203 KSEVDIYERFR-----LYC-----LGKLR-----SNPW----- 225  
Db 324 KIKIDYKRCRGKDKEDKDRYCSRNGYDEKTKRAIGKRGKOCISCLYACNPFYDWIN 383  
QY 226 SELDGLQPETKLINEOLEKINTKGFLLTINSOPAVNGEKSD---SPTVWGCGPGGYVQKA 282  
Db 384 NKEQFQKOKKKYDEIKKYE-----NGASGSGRQKRDAGGTTTNYDG-----YEKK 431  
QY 283 Y-----VEFCSEKELDALVDCKDRT--SLITYAVNKGDSKMSNNGQDVNAV 329  
Db 432 FYDELINKSEYRTVDFLEKLSNEICTKYKDEEGGTFDKNVSD---STSGASGTNYE 487  
QY 330 TWGVFPAKEIIOP-----TIYDVSEFNWKMDEAFETWISRGWASLY---PEDEASRKLY 379  
Db 488 SGTGYFRSKYQCPGCPYCGVKKVNNGSSNMEWERNKNGCKSG--KLYEPKPREGTTITI 545  
QY 380 EEYGGSHFLVSLVDN---DYINGD 400  
Db 546 LKSGKHDIIEKLNKFCDEKNGD 569

RESULT 10  
US-09-210-288-12  
; Sequence 12, Application US/09210288  
; Patent No. 6392026  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/210,288  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fuller, Michael  
; REGISTRATION NUMBER: 36,516  
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176

;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2710 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHEICAL: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Plasmodium falciparum  
;; US-09-210-288-12

Query Match 4.5%; Score 97.5; DB 4; Length 2710;  
Best Local Similarity 19.8%; Pred. No. 3.6;  
Matches 88; Conservative 59; Mismatches 142; Indels 155; Gaps 27;

QY 56 PADIMALL-EPIDNEAVKAVYGIHIGTEM---CKLIAHGKTLHLTYLNMKSALAIL 111  
DB 182 PSQICTMLANSPADIGIVAGRDLYGNPOETKQKQLENNLKITI----- 226  
QY 112 MNLGLIESKYSRSLPWRPRANVRVKEDEVRIFW-ANRPKSYISRTIGMDQYPHGWD 170  
DB 227 --FEKIVKINGAARVAGNDPEFFKLRD---WMTANR-----ETV--WKAITCNAMGN 273  
QY 171 S-----CN---PSYGALSDY--QEMPRARADKKLVEEAVPL----- 202  
DB 274 TYFNATCRGERTKGYCRMDQVPTY---FDIVQYLR-----WFEWAEDFCCKKN 323  
QY 203 KSVEDIYERFR-----LYC-----LGKLR-----SNFW----- 225  
DB 324 KRIVDVNRNGKDKEDKDRYCSRNGYDEKTKRAIGKIRYKQKOCISLYACNYYVDMIN 383  
QY 226 SELGLOPETRIINEOLEKINTGFLTINSOPAVNGEKSD---SPTVWGSGPGGYVOKA 282  
DB 384 NQKQFPKQKKRYDEIKKYE-----NGASGSRQKRDAGTITTYDQ-----YEKK 431  
QY 283 Y-----VEFCSKEKLDALYDKCKDRT--SLTYMAVNDGSKSVNGQTDVNAV 329  
DB 432 FYDELINSEYRTVDKFLKLSNEICTKYKDEGGTIDKKNNSD---STSGASGTNVE 487  
QY 330 TWGVFPKKEIIOF-----TIVDPVSFWMKDEAFELISRCWASLY---PEDEASRKL 379  
DB 488 SQGTFYRSKCYQPCPYGKAVKVNNGSSNMEWEEKNNKCKSG--KLYPRKDKGTTITI 545  
QY 380 EGVGSHFLVSLVDN---DYINGD 400  
DB 546 LKSGGHDIEEKLKFCDEKNGD 569

RESULT 11  
US-08-487-826B-14  
; Sequence 14, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/487,826B  
;; FILING DATE: 10-SEP-1993  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Israelsen, Ned  
;; REGISTRATION NUMBER: 29,655  
;; REFERENCE/DOCKET NUMBER: NIH121.001CPI  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 235-8550  
;; TELEFAX: (619) 235-0176  
;; INFORMATION FOR SEQ ID NO: 14:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3060 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-487-826B-14

Query Match 4.5%; Score 97.5; DB 2; Length 3060;  
Best Local Similarity 19.8%; Pred. No. 4.3;  
Matches 88; Conservative 59; Mismatches 142; Indels 155; Gaps 27;

QY 56 PADIMALL-EPIDNEAVKAVYGIHIGTEM---CKLIAHGKTLHLTYLNMKSALAIL 111  
DB 180 PSQICTMLANSPADIGIVAGRDLYGNPOETKQKQLENNLKITI----- 224  
QY 112 MNLGLIESKYSRSLPWRPRANVRVKEDEVRIFW-ANRPKSYISRTIGMDQYPHGWD 170  
DB 225 --FEKIVKINGAARVAGNDPEFFKLRD---WMTANR-----ETV--WKAITCNAMGN 271  
QY 171 S-----CN---PSYGALSDY--QEMPRARADKKLVEEAVPL----- 202  
DB 272 TYFNATCRGERTKGYCRMDQVPTY---FDIVQYLR-----WFEWAEDFCCKKN 321  
QY 203 KSVEDIYERFR-----LYC-----LGKLR-----SNFW----- 225  
DB 322 KRIVDVNRNGKDKEDKDRYCSRNGYDEKTKRAIGKIRYKQKOCISLYACNYYVDMIN 381  
QY 226 SELGLOPETRIINEOLEKINTGFLTINSOPAVNGEKSD---SPTVWGSGPGGYVOKA 282  
DB 382 NQKQFPKQKKRYDEIKKYE-----NGASGSRQKRDAGTITTYDQ-----YEKK 429  
QY 283 Y-----VEFCSKEKLDALYDKCKDRT--SLTYMAVNDGSKSVNGQTDVNAV 329  
DB 430 FYDELINSEYRTVDKFLKLSNEICTKYKDEGGTIDKKNNSD---STSGASGTNVE 485  
QY 330 TWGVFPKKEIIOF-----TIVDPVSFWMKDEAFELISRCWASLY---PEDEASRKL 379  
DB 486 SQGTFYRSKCYQPCPYGKAVKVNNGSSNMEWEEKNNKCKSG--KLYPRKDKGTTITI 543  
QY 380 EGVGSHFLVSLVDN---DYINGD 400  
DB 544 LKSGGHDIEEKLKFCDEKNGD 567

RESULT 12  
US-08-646-695-2  
; Sequence 2, Application US/08646695  
; Patent No. 6168943  
; GENERAL INFORMATION:  
; APPLICANT: Rose, John K.  
; APPLICANT: RECOMBINANT VESICULOVIRUSES AND THEIR  
; TITLE OF INVENTION: US  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York

COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,695  
FILING DATE: On Even Date Herewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 6523-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 422 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-695-2

Query Match 4.4%; Score 97; DB 4; Length 422;  
Best Local Similarity 19.7%; Pred. No. 0.22;  
Matches 63; Conservative 43; Mismatches 90; Indels 124; Gaps 16;

QY 127 PRRPANEVRKEDYRPIFWANRPK-----YISRTIGMDQYPHGRWGDSCNPS----- 175  
DB 24 PVEYPADYFRKKEI-PLY-INTTKSLDLRGVYQGL-----KSGNVSIIHVN 70  
QY 176 ---YGALSDYQFMPPRRARDKLVEMA-----VPLKSEDIYE----- 210  
DB 71 SYLYGALKDIR-----GKLDKDWSSFGINIGKAGDTIGIFDLVSLKALDGLVLDGVS 122  
QY 211 -----RFRLYCLGKLRSPWSE--LDGLQPEYKIINEOLEKINTKGFLLT 253  
DB 123 DASRTSADDKWLPLYLLGLYRGVGRTOPEYRKKLMDGLTNQCKMINEOFELVPEG---- 178  
QY 254 NSQPAVNGEKSDSPYWGMBGPGYVYOKAYVEF-----CSKEKLDALVYCKDRTS 305  
DB 179 -----RDIFDY-WGNDSENTKIYAADVMEFHKKEHCASFYGTIVSRKDCAA 227  
QY 306 LTYMAVNVKDSWKSINVQOTDVNAVTV-----GVFPAKEI-----IQPTIVD-- 346  
DB 228 LATF-----GHLCKITGMSTEDVTWILNREVADENYQMLPGQELDKADSYMPYLIDFG 282  
QY 347 ---PVSEFNWKDEAFETMSR 363  
DB 283 LSKSPYSYVKNPAPHFHWQ 302

RESULT 13  
PCT-US96-06053-2  
Sequence 2, Application PC/TUS9606053  
GENERAL INFORMATION:  
APPLICANT: Yale University  
TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR  
TITLE OF INVENTION: USES  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06053  
FILING DATE: 01-MAY-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 6523-009-228  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 422 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-06053-2

Query Match 4.4%; Score 97; DB 5; Length 422;  
Best Local Similarity 19.7%; Pred. No. 0.22;  
Matches 63; Conservative 43; Mismatches 90; Indels 124; Gaps 16;

QY 127 PRRPANEVRKEDYRPIFWANRPK-----YISRTIGMDQYPHGRWGDSCNPS----- 175  
DB 24 PVEYPADYFRKKEI-PLY-INTTKSLDLRGVYQGL-----KSGNVSIIHVN 70  
QY 176 ---YGALSDYQFMPPRRARDKLVEMA-----VPLKSEDIYE----- 210  
DB 71 SYLYGALKDIR-----GKLDKDWSSFGINIGKAGDTIGIFDLVSLKALDGLVLDGVS 122  
QY 211 -----RFRLYCLGKLRSPWSE--LDGLQPEYKIINEOLEKINTKGFLLT 253  
DB 123 DASRTSADDKWLPLYLLGLYRGVGRTOPEYRKKLMDGLTNQCKMINEOFELVPEG---- 178  
QY 254 NSQPAVNGEKSDSPYWGMBGPGYVYOKAYVEF-----CSKEKLDALVYCKDRTS 305  
DB 179 -----RDIFDY-WGNDSENTKIYAADVMEFHKKEHCASFYGTIVSRKDCAA 227  
QY 306 LTYMAVNVKDSWKSINVQOTDVNAVTV-----GVFPAKEI-----IQPTIVD-- 346  
DB 228 LATF-----GHLCKITGMSTEDVTWILNREVADENYQMLPGQELDKADSYMPYLIDFG 282  
QY 347 ---PVSEFNWKDEAFETMSR 363  
DB 283 LSKSPYSYVKNPAPHFHWQ 302

RESULT 14  
US-08-965-762-23  
Sequence 23, Application US/08965762  
Patent No. 6280963  
GENERAL INFORMATION:  
APPLICANT: Kollip, Vijal  
TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE  
FILE REFERENCE: 0734/062001  
CURRENT APPLICATION NUMBER: US/08/965,762  
CURRENT FILING DATE: 1997-11-07  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 23  
LENGTH: 347  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-08-965-762-23

Query Match 4.2%; Score 91; DB 4; Length 347;  
Best Local Similarity 21.3%; Pred. No. 0.65;



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2003, 20:30:11 ; Search time 81 Seconds

(without alignments)  
1037.867 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187  
Sequence: 1 ARVYVQTFYDIDLEKFN.....VELVDNDYINGDLFAVEADF 408

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: SPREMBL.21.\*  
2: sp\_archaea.\*  
3: sp\_bacteria.\*  
4: sp\_fungi.\*  
5: sp\_human.\*  
6: sp\_invertebrate.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_protist.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1834	83.9	594	10 Q9SE59	Q9SE59 arabidopsis
2	1834	83.9	594	10 Q947R1	Q947R1 arabidopsis
3	1831	83.7	594	10 Q9SU00	Q9SU00 arabidopsis
4	1777	81.3	592	10 Q9SE60	Q9SE60 arabidopsis
5	1762	80.6	593	10 Q9SE94	Q9SE94 zea mays (m
6	246	11.2	276	16 Q8Y389	Q8Y389 raldstonia s
7	222	10.2	290	16 Q91687	Q91687 pseudomonas
8	221	10.1	314	16 Q8A6F4	Q8A6F4 caudobacter
9	215	9.8	275	16 Q9PEA7	Q9PEA7 xylella fas
10	198	9.1	313	16 Q92NKL	Q92NKL rhizobium m
11	197.5	9.0	163	3 Q94090	Q94090 saccharomyc
12	196	9.0	295	16 Q9CP31	Q9CP31 pasteurella
13	193	8.8	307	16 Q9S2V3	Q9S2V3 streptomyc
14	189	8.6	296	16 Q82302	Q82302 salmonella
15	189	8.6	296	16 Q8X766	Q8X766 escherichia
16	188	8.6	282	16 Q9PN93	Q9PN93 campylobact

17	186	8.5	293	2 Q9L5C1	Q9L5C1 leptospira
18	184	8.4	183	6 Q9NOY5	Q9NOY5 sus scrofa
19	182	8.3	292	16 Q9J202	Q9J202 neisseria m
20	182	8.3	292	16 Q9J207	Q9J207 neisseria m
21	182	8.3	294	2 Q93ER8	Q93ER8 photobacter
22	179	8.2	294	16 Q8ZJ16	Q8ZJ16 yersinia pe
23	179	8.2	312	16 Q8UDK2	Q8UDK2 agrobacteri
24	171	7.8	304	16 Q9KNP6	Q9KNP6 vibrio chol
25	163	7.5	283	16 Q9CG56	Q9CG56 lactococcus
26	162.5	7.4	305	16 Q9K877	Q9K877 rhizobium l
27	162	7.4	274	16 Q8Y186	Q8Y186 bruceella me
28	155.5	7.1	349	5 Q9V710	Q9V710 streptococ
29	154.5	7.1	130	2 Q9ACM5	Q9ACM5 streptococ
30	148.5	6.8	288	16 Q9T530	Q9T530 streptococ
31	142	6.5	318	2 Q9F354	Q9F354 rhodotherm
32	138	6.3	305	2 Q938W5	Q938W5 streptomyc
33	108	4.9	674	17 Q8TJ17	Q8TJ17 methanosarc
34	104.5	4.8	1138	16 Q9PK07	Q9PK07 chlamydia m
35	104	4.8	483	12 Q9ODJ3	Q9ODJ3 clover yell
36	103	4.7	538	12 Q67458	Q67458 figwort mos
37	102	4.7	706	5 Q9NDE9	Q9NDE9 dictyostell
38	102	4.7	920	13 Q910A4	Q910A4 brachydanio
39	101	4.6	301	16 Q8R926	Q8R926 thermotanaer
40	101	4.6	548	3 Q9P8J5	Q9P8J5 candida alb
41	100	4.6	422	12 Q88993	Q88993 vesicular s
42	100	4.6	1142	16 Q97J24	Q97J24 clostridium
43	99.5	4.5	619	5 Q9VMI9	Q9VMI9 dirosophila
44	99	4.5	406	12 Q83136	Q83136 bean yellow
45	98.5	4.5	582	16 Q67141	Q67141 aquilex aeo

## ALIGNMENTS

RESULT 1  
Q9SE59 PRELIMINARY: PRT; 594 AA.  
ID Q9SE59  
AC Q9SE59;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Methylenetetrahydrofolate reductase MTHFR2 (EC 1.5.1.20)  
DE (At2g44160/F6E13.29).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20062814; PubMed=10593891;  
RA Koje S., Wang H., McNeil S.D., Raymond R.K., Appling D.R.,  
RA Snachar-Hill Y., Bohnert H.J., Hanson A.D.;  
RT "Isolation, characterization and functional expression of cDNAs  
RT encoding NADH-dependent methyltetrahydrofolate reductase from  
RT higher plants.";  
RL J. Biol. Chem. 274:36089-36096(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Shih P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,  
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,  
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamitani A.,  
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,  
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,  
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,  
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RT "Arabidopsis cDNA clones";  
RT Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF181967; AAD5788.1; -;  
DR EMBL; AY050434; AAK91450.1; -;  
DR HSSP; P00394; 1B5T.  
DR InterPro; IPR000194; ATPase\_a/bcentre.



[illegible]

#### RESULT 4

ID	09560	PRELIMINARY;	PRT;	592 AA.
AC	09560;			
DT	01-MAY-2000 (TrEMBLrel. 13)			Created
DT	01-MAY-2000 (TrEMBLrel. 13)			Last sequence update
DT	01-JUN-2002 (TrEMBLrel. 21)			Last annotation update
DE	Methyltetrahydrofolate reductase MTHFR1 (EC 1.5.1.20) (Putative			
DE	methyltetrahydrofolate reductase MTHFR1).			
GN	F24616.240 OR AT3539970.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	{1}			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=20062814; PubMed=10593891;			
RX	Roje S., Wang H., McNeil S.D., Raymond R.K., Applng D.R.,			
RA	Shachar-Hill Y., Bohmert H.J., Hanson A.D.;			
RT	Isolation, characterization and functional expression of cDNAs			
RT	encoding NADH-dependent methyltetrahydrofolate reductase from			
RT	higher plants. "			
RL	J. Biol. Chem. 274:36089-36096(1999).			
RN	{2}			
RN	SEQUENCE FROM N.A.			
RP	D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W.,			
RA	Lamcke K., Meyer R.F.X., Quetier F., Salanoubat M.;			
RL	Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.			
RN	{3}			
RP	SEQUENCE FROM N.A.			
RP	BY Arabidopsis sequencing project;			
RL	Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.			
RN	{4}			
RN	SEQUENCE FROM N.A.			
RA	Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,			
RA	Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,			
RA	Yamamura Y., Yu G., Yu S., Bowser L., Gaminchi P., Chen H., Cheuk R.			
RA	Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,			
RA	Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,			
RA	Narusaka M., Nguyen M., Palm C.J., Sakurai T., Setou M., Seki M.,			
RA	Shin P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,			
RA	Theologis A.;			
RT	Full length cDNA of gene At3559970 (GI:15232215). "			
RL	Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AP01966; AA05787.1;			
DR	EMBL; AL18647; CA075816.1;			
DR	EMBL; AF070034; AA149791.1;			
DR	HSSP; P00394; 1B5T;			
DR	InterPro; IPR000194; ATrase_a/centre.			

DR InterPro: IPR004621; Fadh2\_euk.  
DR InterPro: IPR003171; Mchydrof\_redtase.  
DR Pfam: PF02219; MTHFR\_1.  
DR TIGRfams: TIGRf00677; Fadh2\_euk\_1.  
DR PROSITE: PS00152; ATPase\_ALPHA\_BETA; UNKNOWN\_1.  
KO Oxidoreductase.  
KO SSIDNUCE 592 AA; 66288 MW; 72D7453AF1A1573 CRC64;

Query Match	81.3%;	Score 1777;	DB 10;	Length 592;
Best Local Similarity	78.6%;	Pred. No. 1.2e-143;		
Matches 319;	Conservative 44;	Mismatches 43;	Indels 0;	Gaps 0;

Oy	1	ARVTVQOLEYDDIDFLAKVNDNCROIGITCIVGIMPIINNYGFIIRMTGEFKTIPADII	60
Db	187	ADLIIVQJFYDDIDFLKRYVNCROIGINCPIVGMPIISNYGFLRMAGFECKTIRPAELI	246
Oy	61	AALEPIKONDEAVKAYGITHLGTECKKILAHGKITLHLYTLMMESALATIIMNLGLIES	120
Db	247	AALEPIKONDEAVKAYGITHFATECKKILAHGTSITLHLYTLVNDKSAIGILIMNLGLIDES	306
Oy	121	KVSSLEPWRPRANVRVEDEVPJPEFMANRBRKSYISRTITMDQYPHGRMGSCNPSYGALS	180
Db	307	KISLSLEPWRPRANVRTEKEDVPRJEFMANRBRKSYISRTKCMNDFPHGRMGDSHAASYTIS	366
Oy	181	DYQPMRPRAODKKLYBEWNAVPLKASVEDIYERFPLVCLGLKRSNPMSELDGLQPTTKIINE	240
Db	367	DYQPMRPRAODKKLYQEWNAVPLKASIEVOQKEFELCTIGMLKSSPMSELDGLQPTTKIINE	426
Oy	241	QLEKINTKGFILITINSOPAVNGEKSDSPTEVMGGPGGYVYOKKAYVEFECSREKIDALVDC	300
Db	427	OLGINSNGELITINSOPASNAKADSPAIMGGPGGYVYOKAYLEFECSKDKLDTIYEXS	486
Oy	301	KDRSLTYANVANKDSMKSNNYGOFDVAAYVMAGVPAKKEIIQTYIYDPSFVNMWDEAFEL	360
Db	487	KAFPSITTYAAVKKSEMWVSNJNGEDVAAVYMGVPAKEVIQTYIYDPSFVNMWDEAFEL	546
Oy	361	WSRGMSLYPDEDDASRKLYEEVYGSHFLVSLVNDYINDGLFANVA	406
Db	547	WSRSMALYLPEDDPSRKILLEYKNSYSLVLYNNYINDIDFSVA	592

## RESULT

ID	NAME	PROTEIN	SEQUENCE
AC	Q9SE94	PRELIMINARY	593 AA.
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)	
DE	Methyltetrahydrofolate reductase (EC 1.5.1.20).		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC Clade		
OX	Panicoideae; Andropogonaceae; Zea.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-20062814; PubMed-10593891;		
RA	Roje S., Wang H., McNeill S.D., Raymond R.K., Applng D.R.,		
RA	Shachar-Hill Y., Bohnett H.J., Hanson A.D.;		
RT	"Isolation, characterization and functional expression of cDNAs		
RT	encoding NADH-dependent methyltetrahydrofolate reductase from		
RT	higher plants."		
RL	J. Biol. Chem. 274:36089-36096(1999).		
DR	EMBL: AF174486; AAD51733.1; -.		
DR	HSSP: P00394; 1B5F.		
DR	InterPro: IPR000194; ATPase_a/bcentre.		
DR	InterPro: IPR004621; Fadh2_euk.		
DR	InterPro: IPR003171; Methyltrf_reductse.		
DR	Pfam: PF02219; MTHFR.1		
DR	TIGRFAMs: TIGR00677; fadh2_euk; 1.		
DR	PROSITE: PS00152; ATPase_A/PPA_BETA.		
DR	UNKNOWN.1.		
DR	Sequence.		
SC	SEQUENCE	593 AA;	45689675B638B4 CRC64;

Query Match	80.6%;	Score 1762;	DB 10;	Length 593;
Best Local Similarity	80.0%;	Pred. NO. 2.4e-142;		
Matches 324;	Conservative 35;	Mismatches 46;	Indels 0;	Gaps 0;

QY	1	AAVAVTQAFVDIIOFLFKFVNOCROIGTICIVGIMPIINNKKGIIRNTGCGCKTILPAIDM	60
Db	187	ADLIYVQAFVDIIOFLFKFVNOCROIGTICIVGIMPIINNKKGIIRNTGCGCKTILPAIDM	246
QY	61	AALEPIYDNEEAKVKAIGHILGTEWCKKILAHGKTLHTLTATMKSALATIANMLGLEES	120
Db	247	AAALDIKDNBEAAVROGIIHILGTEWCKKILATGKTLHTLTATMKSALATIANMLGLEES	306
QY	121	KVSRSLPWRPAAVPEVKEDVAPLPMANRKRXYISITIGMDYDPRGMSDSCNPSYGLS	180
Db	307	KYSRSLPWRPAAVPEVKEDVAPLPMANRKRSTYLKATGMQYDPRGMSDSCNPSYGLT	366
QY	181	DYQFRRPARDKILVEEMAVPLKSEVDIYRPFYLQGLKLSNWSRSLDIOPEKRIINE	240
Db	367	DHQFLRRPGRGSKLQGEEMAVPLKSEVDISRFLTNFQCKLTSRWSRSLDIOPEKRIIIO	426
QY	241	OLEKLNKIGELITNSOPAAVNGEKSDPFWGMSGPGSIYYQKAYVEFFCSKEKLDALVDC	300
Db	427	OLVNTINQGEFLITNSOPAAVNGEKSDSPFWGMSGPGSIYYQKAYLEFFCAEKELDOLIEKI	486
QY	301	KDRTSILYMAVNNDSQSMKSNVQOTDVAAYVWGVFPKKEILOPTIVDSVFNWKKCEAEI	360
Db	487	KAFPSLTVIYAVNNDSGFENISPNANVAATWGVFPKKEILOPTIVDAHSPMWKDBAEI	546
QY	361	NSRGHSLIYPEDDASRKLYEEVGSHTLVSLVNDYINGDLFANF	405
Db	547	WTRGMCNFBPGDSSRELEKQATYVLVSLVNDYVGGDLFAAF	591

RESULT 6	
Q8Y389	
ID Q8Y389	PRELIMINARY;
276 AA	PRT;

DT	01-MAR-2002	(TREMBLrel. 20, Created)
DT	01-MAR-2002	(TREMBLrel. 20, Last sequence update)
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)
DE	Probable 5,10-methylerythritohydrofolate reductase oxidoreductase	
DE	protein (EC 1.7.99.5)	
GN	MEFPO OR RSC0091 OR RSO2264.	
OS	Ralstonia solanaceae (Pseudomonas solanacearum).	
OC	Bacteria; Proteobacteria; beta subdivision; Ralstonia group;	
OC	Ralstonia.	
OX	NCBI_TaxID=305;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CM1100;	
RC	MEDLINE=21681879; PubMed=11823852;	
RA	Salanoubat M., Genin S., Artiguenave F., Gonzy J., Mangenot S.,	
RA	Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,	
RA	Chandler M., Christine N., Glaudel-Renard C., Guinac S., Denange N.	
RA	Gaspin C., Layle M., Moisan A., Robert C., Saurin W., Schlex T.,	
RA	Siguler P., Thebaud P., Whalen M., Wincker P., Levy M.,	
RA	Weissenbach J., Bouchet C.A.;	
RT	"Genome sequence of the plant pathogen Ralstonia solanacearum."	
RL	Nature 415:497-502(2002).	
DR	EMBL: AL646057; CAD13619.1; -.	
DR	InterPro: IPR004620; Fadn2; bact.	
DR	InterPro: IPR003171; Methylrof_redctase.	
DR	Pfam: PF002319; MHRF_1	
DR	TIGRFAMS; TIGR00676; Fadn2; 1.	
DR	Oxidoreductase; Complete proteome.	
QO	SEQUENCE 2/6 AA: 30770 MW: 3A5DC860DB63A2B CRC64;	

Query Match	11.28;	Score 246;	DB 16;	Length 276;
Best Local Similarity	40.58;	Pred. No. 6.4e-13;		
Matches 47;	Conservative 23;	Mismatches 46;	Indels 0;	Gaps 0;

QY 1 ARVIVTQLEFYDTIDIFLKFNDCRQIGITCPIVPGIMPINNYKGFIRM TGFCCKTIPADIM 600

Db	161	ANSAITQFFFNADAYQFQVDDARKLGTADVIVIGIMITNSOALRPSRECKGATLPIWIA	220
Oy	61	AALEPTKQNEAVKAVGILHCEMKCKTLIAHGKTLHELYTNMEKSLATLAINMLG	116
Db	221	KRLSEFGSDRSIRACGLDVAYNALCDRLILAGCAPGLHFTTLNAAAARAIKQRL	276
RESULT 7			
ID	Q91687	PRELIMINARY;	PRT; 290 AA.
AC	Q91687;		
DT	01-MAR-2001 (TREMBrel, 16, Created)		
DT	01-MAR-2001 (TREMBrel, 16, Last sequence update)		
DT	01-JUN-2002 (TREMBrel, 21, Last annotation update)		
DE	5,10-methyltetrahydrofolate reductase.		
GN	MEP or PA0430		
OS	Pseudomonas aeruginosa.		
CC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;		
CC	Pseudomonas		
NC	NCBI_TaxID=287;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC 15692 / PA01;		
RX	MEDLINE=20437337; PubMed=10984043;		
RA	Stover C.R., Pham X.-O.T., Ervin A.L., Mizoguchi S.D., Warren P.,		
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,		
RA	Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,		
RA	Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,		
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Poulsen I.T.,		
RA	Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;		
RT	"Complete genomic sequence of Pseudomonas aeruginosa PA01, an		
RT	opportunistic pathogen.";		
RL	EMBL; AE004480; AAC03819.1; -		
DR	HSSP; P00394; 1B5T.		
DR	InterPro; IPR004620; Fadh2_bact.		
DR	InterPro; IPR003171; Methylcrof_reductase.		
DR	Pfam; PF02219; TTHFR_1.		
DR	TIGRFAMs; TIGR00676; fadh2_1.		
DR	Complete proteome.		
QW	SSQID= 290 AA. 39228 MW; 39256FA43D896DA CRC64;		

Query Match	10.28;	Score 222;	DB 16;	Length 290;
Best Local Similarity	38.48;	Pred. No. 7.8e-11;		
Matches 43;	Conservative 22;	Mismatches 47;	Indels 0;	Gaps 0;

[illegible]

RESULT 8		
09A6F4		
ID	09A6F4	PRELIMINARY; PRT; 314 AA.
AC	09A6F4;	
DT	01-JUN-2001 (TrEMBLrel. 17, Created)	
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
DE	5,10-methylmenetetrahydrofolate reductase.	
GN	CC2140.	
OS	Caulobacter crescentus.	
OC	Bacteria; Proteobacteria; alpha subdivision; Caulobacter group	
CC	Caulobacter.	
OX	NCBI_TaxID=155892;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 19089 / CB15;	
XX	MEDLINE=21173698; PubMed=11259647;	

RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Padlock N.D., Ely B.,  
 RA DeJoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,  
 RA Kolonay J.F., Sait J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Toljaczek T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,  
 RT "Complete genome sequence of *Caulobacter crescentus*,"  
 RA Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 RL EMBL: AE005886; AAK24111.1; -.  
 DR HSSP: P00394; 1B5T.  
 DR TIGR: CC2140; -.  
 DR InterPro: IPR004620; Fadh2\_bact.  
 DR InterPro: IPR003171; Methylrof\_redctse.  
 DR Pfam: PF02219; MTHFR; 1.  
 DR TIGRFAMS: TIGR00676; fadh2; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 314 AA; 34247 MW; 521E7A352E34A9C5 CRC64;

Query Match 10.1%; Score 221; DB 16; Length 314;  
 Best Local Similarity 37.9%; Pred. No. 1;e-10;  
 Matches 44; Conservative 21; Mismatches 51; Indels 0; Gaps 0;

OY 5 VTQLFYDTDFLKFVNDRCRGITGTCPIYVGIMPINNYKGFTRMTGFCCKTKIPADIMALE 64  
 DB 191 ISOFFDIDALRFYDKYRAGITIPYVGPIMPTNFAGLKKMAACQTAIPSWGLFD 250  
 OY 65 PIKONEEAVKAYGTHLGTETMCKKILAHGKTLHLTYLNMESALAIIMNLG 120  
 DB 251 GLENDATRRILACSVAAEMCKALOEQGFDPHFYTLRADLVYALICRVGVRIS 306

## RESULT 9

O9PEA7 PRELIMINARY; PRT; 275 AA.  
 AC O9PEA7;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DE 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE 5,10-methylenetetrahydrofolate reductase.  
 GN Xf1121.  
 OS *Xyella fastidiosa*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Rehnach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carriro D.M., Carrier H.,  
 RA Colauro N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
 RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gubler A.,  
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kurame E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorallo C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A. Jr., Norega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasak H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatr M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE003948; AAF83931.1; -.  
 DR HSSP: P00394; 1B5T.  
 DR InterPro: IPR003171; Methylrof\_redctse.  
 DR Pfam: PF02219; MTHFR; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 275 AA; 30499 MW; 22D6C156B53241D0 CRC64;

Query Match 9.8%; Score 215; DB 16; Length 275;  
 Best Local Similarity 33.9%; Pred. No. 2;9e-10;  
 Matches 39; Conservative 30; Mismatches 46; Indels 0; Gaps 0;

OY 1 ARVITQLFYDTDFLKFVNDRCRGITGTCPIYVGIMPINNYKGFTRMTGFCCKTKIPADIM 60  
 DB 160 ADAITQYFNPDPAYPHFVDAVORLGVTLPIYAGVDPISNFKLHHSQCGAETPRWT 219  
 OY 61 AALEPIKONEEAVKAYGTHLGTETMCKKILAHGKTLHLTYLNMESALAIIMNLG 115  
 DB 220 KKNQATGDDIKTSIRAGADVYTAALCERLLAGAPGLHFTYTLNARFSTGYLRIG 274

## RESULT 10

O92NKL PRELIMINARY; PRT; 313 AA.

AC O92NKL;  
 DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DE 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Probable 5,10-methylenetetrahydrofolate reductase oxidoreductase  
 DE protein (EC 1.7.99.5).  
 GN METE OR R02199 OR SMC01843.  
 GN Rhizobium meliloti (Sinorhizobium meliloti).  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gonzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaune V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsberger U.,  
 RA Renard C., Thebaud P., Vandendol M., Weidner S., Gallibert F.,  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 Sinorhizobium meliloti strain 1021."  
 RT Sinorhizobium meliloti strain 1021.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL: AL591789; CAC46778.1; -.  
 DR InterPro: IPR004620; Fadh2\_bact.  
 DR InterPro: IPR003171; Methylrof\_redctse.  
 DR Pfam: PF02219; MTHFR; 1.  
 DR TIGRFAMS: TIGR00676; fadh2; 1.  
 KM Oxidoreductase; Complete proteome.  
 SQ SEQUENCE 313 AA; 34196 MW; C8BR3136B0D15071 CRC64;

Query Match 9.1%; Score 198; DB 16; Length 313;  
 Best Local Similarity 34.2%; Pred. No. 1;e-08;  
 Matches 38; Conservative 20; Mismatches 53; Indels 0; Gaps 0;

OY 5 VTQLFYDTDFLKFVNDRCRGITGTCPIYVGIMPINNYKGFTRMTGFCCKTKIPADIMALE 64  
 DB 180 LTOFFPNDFFEXYLEVNRAGIAITVPEIILPINMLTOYQKAGICGARGVESVSRIG 239  
 OY 65 PIKONEEAVKAYGTHLGTETMCKKILAHGKTLHLTYLNMESALAIIMNLG 115  
 DB 240 PIDDQPEREREAHAIAAEQVADLVRRGVSDFLTYMNRQLTAVCELMG 290

## RESULT 11

O94090



```

RESULT 14
082302 ID 082302 PRELIMINARY; PRT; 296 AA.
AC 082302;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase.
GN str3761.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
CX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Holden M.T.G., Sebatina M.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Kirogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AL627279; CAB09516.1; -
DR InterPro; IPR004620; Fadh2_bact.
DR InterPro; IPR003171; Methylrof_redctse.
DR Pfam; PF02219; MTHFR; 1.
DR TIGRFAMs; TIGR00676; fadh2; 1.
KW Complete proteome.
SQ SEQUENCE 296 AA; 33202 MW; 3165CE17678ACBEE CRC64;

Query Match 8.6%; Score 189; DB 16; Length 296;
Best Local Similarity 32.1%; Pred. No. 5.4e-08;
Matches 36; Conservative 25; Mismatches 51; Indels 0; Gaps 0;

QY 5 VTOLFVDTDFLEKFNVDRCQIGITCPVIGIMPINNYKGFIRMTGFCCKTIPADIMALE 64
DB 181 ITQFFDVESYLRFRDRCVSAGIDVELIPGLPVSFKQAKKPADMTNVRIRPMSLSLME 240

QY 65 PIRKNEAVKAYGIHLGTEMCKKIIAHGIKTLHLTYTLNMEKSALAILMNLGL 116
DB 241 GDDDAETRKLVGANIAMDMVKILISREGVDFEFTYTLNRAEMSYAICHTLGV 292

RESULT 15
08X766 ID 08X766 PRELIMINARY; PRT; 296 AA.
AC 08X766;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase.
GN MEMP OR Z5496 OR ECS4870.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
CX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postfal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;

```

```

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005625; AAG59142.1; -
DR InterPro; IPR002567; BAB38293.1; -
DR InterPro; IPR004620; Fadh2_bact.
DR InterPro; IPR003171; Methylrof_redctse.
DR Pfam; PF02219; MTHFR; 1.
DR TIGRFAMs; TIGR00676; fadh2; 1.
KW Complete proteome.
SQ SEQUENCE 296 AA; 33129 MW; 1FA66221D8296676 CRC64;

Query Match 8.6%; Score 189; DB 16; Length 296;
Best Local Similarity 32.1%; Pred. No. 5.4e-08;
Matches 36; Conservative 25; Mismatches 51; Indels 0; Gaps 0;

QY 5 VTOLFVDTDFLEKFNVDRCQIGITCPVIGIMPINNYKGFIRMTGFCCKTIPADIMALE 64
DB 181 ITQFFDVESYLRFRDRCVSAGIDVELIPGLPVSFKQAKKPADMTNVRIRPMSLSLME 240

QY 65 PIRKNEAVKAYGIHLGTEMCKKIIAHGIKTLHLTYTLNMEKSALAILMNLGL 116
DB 241 GDDDAETRKLVGANIAMDMVKILISREGVDFEFTYTLNRAEMSYAICHTLGV 292

Search completed: February 14, 2003, 23:09:55
Job time : 86 secs

```

1  
2  
3

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2003, 23:08:27 ; Search time 2159 Seconds  
(without alignments)  
3060.564 Million cell updates/sec

Title: US-09-720-451-6  
Perfect score: 2187  
Sequence: 1 ARVLTQLEFYDIDFLKFN.....VSLVNDYINGDLFAVAFADF 408

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=x1h  
-O=/cgn2.1/USPFO/US09720451/runat\_05022003\_074816\_2604/app.query.fasta.1.583  
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09720451\_@cgn.1.1\_763\_@runat\_05022003\_074816\_2604 -ICPU=6 -ICPU=3  
-NARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pto:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1750	80.0	2423	11	AY109096	AY109096 Zea mays
2	1400	64.0	851	14	BQ165347	BQ165347 EST611216
3	1216	55.6	759	13	B1308605	B1308605 EST530015
4	1129	51.6	934	12	BG365861	BG365861 HVSME1000
5	1128	51.6	823	12	BC418616	BC418616 HVSMEK002
6	1104	50.5	666	10	AW201065	AW201065 se97d12.y
7	1097	50.2	760	14	BO994876	BO994876 OGPF821.y
8	1093	50.0	621	10	AM830074	AM830074 sm22h02.y
9	1087	49.7	700	10	BE576981	BE576981 LO-1586T3
10	1079.5	49.4	752	13	BI421654	BI421654 EST532320
11	1052.5	48.1	842	12	BC418214	BC418214 HVSMEK002
12	1052	48.0	611	12	BER05904	BER05904 ss62h07.y
13	1049	48.0	729	12	BC123885	BC123885 EST469531
14	1045	47.8	821	10	BE643238	BE643238 Cr12_8_E1
15	1042	47.6	588	12	BF071278	BF071278 st42e03.y
16	1029	47.1	755	13	BI308302	BI308302 EST529712
17	1021	46.5	911	10	BE231181	BE231181 HVSMEK001
18	1017.5	46.5	836	12	BC418185	BC418185 HVSMEK002
19	1012	46.3	573	10	AW203249	AW203249 sf27g11.y
20	1011	46.2	579	10	AW203207	AW203207 sf27d02.y
21	1011	46.2	652	12	BG904512	BG904512 Talr1132g
22	1008	46.1	668	14	BQ404086	BQ404086 GA_Ed006
23	1004	45.9	568	10	BE346411	BE346411 sp25c05.y
24	996	45.5	573	10	AW703738	AW703738 sk23g06.y
25	995	45.5	569	14	BM885389	BM885389 sal19b11.y
26	980	44.8	357	10	AW61036	AW61036 sl17b07.y
27	980	44.8	593	10	BE204232	BE204232 EST396908
28	976	44.6	627	10	AV915401	AV915401 AV915401
29	966	44.2	557	13	BI425670	BI425670 sah7b0b07.y
30	965	44.1	634	10	BE346389	BE346389 sp25a05.y
31	960	43.9	575	10	AW733729	AW733729 sk77e01.y
32	955	43.7	757	17	BH012892	BH012892 TDGA075TH
33	947	43.3	676	10	BE403890	BE403890 WHED420_H
34	944	43.2	565	9	AI900449	AI900449 sc07a01.y
35	943.5	43.1	721	14	BO993154	BO993154 OGF27118.y
36	936	42.8	550	12	BG041679	BG041679 sv37e08.y
37	936	42.7	661	14	BQ863196	BQ863196 QGC23D15.y
38	934	42.8	814	10	BE194793	BE194793 HVSMEK008
39	927	42.4	567	10	AM584093	AM584093 N210025e
40	917	41.9	582	10	AM931213	AM931213 EST357056
41	911	41.7	562	13	BM528573	BM528573 sal67h06.y
42	899	41.1	566	10	BE405193	BE405193 WHE1211.D
43	893	40.8	556	10	BE433601	BE433601 EST400130
44	893	40.8	558	10	BE433609	BE433609 EST400138
45	888.5	40.6	733	12	BE705725	BE705725 sc02_02d0

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AY109096	AY109096	Zea mays PC0119715 mRNA sequence.	AY109096	AY109096.1	GI:21212506	HTC.	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC	1 (bases 1 to 2423)	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 2423)  
AUTHORS Coe,E.C.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

FEATURES  
source  
1.2423  
/organism="zea mays"  
/db\_xref="MaizeDB:636961"  
/db\_xref="taxon:4577"  
/clone="PC0119715"  
/clone.lib="Maize Mapping Project/DuPont Consensus Library"  
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 630 a 521 c 619 g 648 t 5 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.96e-202 Length: 2423  
Score: 1750.00 Matches: 323  
Percent Similarity: 88.40% Conservatve: 35  
Best Local Similarity: 79.75% Mismatches: 47  
Query Match: 80.02% Indels: 0  
DB: 11 Gaps: 0

US-09-720-451-6 (1-408) x AY109096 (1-2423)

QY 1 AlAaYValIleValIthrGlnLeuPheTyrrAspThrAspIlePheLeuLysPheValAsn 20  
DB 618 GCTGACCTTATGTGTACACAACTTCTATGATACCGACATCTTCTCAAGTTTGTGAAT 677

QY 21 AspCysArgGlnIleGlyIleThrCysProIleValIleProGlyIleMetProIleAsn 40  
DB 678 GACTGGCGCAATGTGTATTAATCTCCCATTTCTCCGGCATATATGCCAATAATAC 737

QY 41 TyrlGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60  
DB 738 TACAAGGTTTCTCGCGACTGACTGGCTCTGCAAACTAAGATACCTTCTGAGATCACT 797

QY 61 AlAlAlaLeuGlnProIleLysAspAsnGlnGluAlaValLysAlaIleGlyIleHisLeu 80  
DB 798 GCTGCACTAGATCTTATCAAAAGCAATGAGAGGCTGTTAGACAATATGAAATCCACCTT 857

QY 81 GlyThrGlnMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrrThr 100  
DB 858 GGAACGAGATGTGCAAGAAATTTCTGCTACTGCGATTAAGACTTTTGACACTTTACACA 917

QY 101 LeuAsnMetGlnLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlnLysLeu 120  
DB 918 CTAAACATGACAGACGCTGCTATAGGAATTTGATGAATCTTGATATATGAGAGGCTC 977

QY 121 LysValSerArgSerLeuProTrrpAlaGArgProAlaAsnValPheArgValLysGlnAsp 140  
DB 978 AAGGTTTCAAGGCCATTACTTGGAGCGCAGCAGCTAATGTTTCCGTGTTAAAGGAGAT 1037

QY 141 ValArgProIlePheThrAlaAsnArgProLysSerTyrrIleSerArgThrIleGlyTrrp 160  
DB 1038 GTTCGACCTATATTTGGGCGCAACACCAAGAGCTATCTTAAAGGACATTAGGTTGG 1097

QY 161 AspGlnTyrrProHisGlyArgTrrpGlyAspSerCysAsnProSerTyrrGlyAlaLeuSer 180  
DB 1098 GATCACTATATCCCATGAGAGGTGGGTGATTTCTCGGAACCATCATATGAGACCTACT 1157

QY 181 AspTyrrGlnPheMetArgProAlaArgAspLysLysLeuValGlnIleTrrpAlaVal 200  
||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 1158 GACCAACAGTTTCAACAAGACCAAGAGCCGCTGTAGAAGCTTCAAGAGAAATGGCTGTT 1217  
QY 201 ProLeuLysSerValGluAspIleTyrrGluArgPheArgLeuTyrrCysLeuGlyLysLeu 220  
DB 1218 CCACGGAATCTGTGGAGACATTAAGTATGAGCGCTTACAAACTTGTGTCAAGGAACTC 1277

QY 221 ArgSerAsnProTrrpSerGlnLeuAspGlyLeuGlnProGlyTrrpLysIleIleAsnGln 240  
DB 1278 ACAAGCACCCATGTGTGCAATTTGAGCGGCTTCAACACAGACAAAGAAATATGATGATC 1337

QY 241 GlnLeuGlnLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260  
DB 1338 CAGTTGGTGAATATTAAACCAAGAGGTTTCTTACAAATTAACACCAACTGCTGTAAAT 1397

QY 261 GlyGlyLysSerAspSerProThrValGlyTrrpGlyArgProGlyTyrrValTyrrGln 280  
DB 1398 GGAGGAATCCGACTCCCTACTGTGTGGGTGGGTGCTGTGAGGCTACGTTATACG 1457

QY 281 LysAlaTyrrValGluPhePheCysSerLysGlnLysLeuAspAlaLeuValAspLysCys 300  
DB 1458 AAGGCTTACCTCGAATTTCTTGGCGCAAGAGAAAGTTGACCAACTATATGAGAAATC 1517

QY 301 LysAspArgThrSerLeuThrTyrrMetAlaValAsnLysAspGlySerTyrrPlysSerAsn 320  
DB 1518 AAGCATTCCTCTCTCACTTACATTCCTGTGAACAAGATGAGAAACATTTCTCAAT 1577

QY 321 ValGlyGlnThrAspValAsnAlaValThrTrrpGlyValPheProAlaLysGlnIleIle 340  
DB 1578 ATTTCACCGAAGCCCGTAAATGCTGTCAAGTGGGTGTGTTTCCCTGCCAAGAGATATC 1637

QY 341 GlnProThrIleValAspProValSerPheAsnValTrrpLysAspGluAlaPheGluIle 360  
DB 1638 CAGCTTACGTTGTGATCATGATCAAGATTATTTGTTGGAAGAGACGAAGCATTTGAATC 1697

QY 361 TrrpSerArgGlyTrrpAlaSerLeuTyrrProLysAspGluAlaSerArgLysLeuValGln 380  
DB 1698 TGGACTCGGGGTGGGTGGATGATCTTCCGTGAGGTATGATGAGGAGCTATGAG 1757

QY 381 GluValGlyGlySerHisPheLeuValSerLeuValAspAsnAspTyrrIleAsnGlyAsp 400  
DB 1758 AAGTTCGAAGACCTACTACTGTCAGCTCTTACAAACGACTACGTCAGGGGAC 1817

QY 401 LeuPheAlaValPhe 405  
DB 1818 CTGTTTGTGCTTC 1832

RESULT 2  
LOCUS B0165347 851 bp mRNA linear EST 25-APR-2002  
DEFINITION EST611216 KYKC Medicago truncatula cDNA clone pKYC-8D7, mRNA sequence.  
ACCESSION B0165347  
VERSION B0165347.1 GI:20307667  
KEYWORDS EST.  
SOURCE  
ORGANISM Medicago truncatula  
barrel medic.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.  
REFERENCE 1 (bases 1 to 851)  
AUTHORS Vanderbosch,K., Endre,G., Silverstein,K., Town,C.D., Van Aken,S., Uiterback,T., Cheung,F. and Fraser,C.M.  
TITLE The Medicago truncatula 'klionone' set: ESTs selected and re-arrayed from various libraries  
JOURNAL Unpublished (2002)  
CONTACT: Vanderbosch K  
Department of Plant Biology  
University of Minnesota  
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA  
TEL: 612 624 2755  
Fax: 612 625 1738  
Email: kvanderb@cbs.umn.edu





```

Db      61 CCTGGAGGCCACCACTAATGTTTCCGTGCAAGAGGATGTCGCCCTATATTCTGG 120
QY      147 AIAASATGProlyserTyrTlleSerArgTlleGlyTTPAspGlyTyrProHsgly 166
Db      121 GCCAACAGACCAAGATTACATTCAGAGACCAGCTGTGGATCAATACCAACAGGGG 180
QY      167 AcGTPGlyAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyrGlnPheMetArg 186
Db      181 CGGTGGGCGATTCCAGAGAACCATCATACGAGACGCTTAATAGCACACAGTTCACACGG 240
QY      187 ProArgAlaArgAspLysLysLeuValGluGluTTPAlaValProLeuLysSerValGlu 206
Db      241 CCACGTGGGCGGTGTGAAGAGCTCCAAAGAGATGGCGCTTCCACATGAAATCTGTGCA 300
QY      207 AspGlyTyrGluArgPheArgLeuTyrCysLeuGlyLysLeuArgSerAsnProTyrSer 226
Db      301 GACATTAATGAGCGGTTCTTAACCTTCTGTAAGGAAAACTTAAGACGCCATGTCT 360
QY      227 GluLeuAspGlyLeuGlnProGluTyrTyrLysLleIleAsnGluGlnLeuGluLysIleAsn 246
Db      361 GAGTTAGATGCTTCAACCCGAGACAGATATGATGATCAGCTGTGAAGATGAAC 420
QY      247 ThrLysGlyPheLeuThrLleAsnSerGlnProAlaValAsnGlyLysLysSerAspSer 266
Db      421 TCAAGGGTTCTTACCATCAACAGCCAGCACTGTGTAATGACAGAAATCTGAGTCT 480
QY      267 ProThrValGlyTyrProGlyLysTyrValTyrGlnLysAlaTyrValGluPhe 286
Db      481 CCAAGTGTGATGGGGCGGTCCANGAGGCTATGTTACAAAAGGCTACGCAATTC 540
QY      287 PheCysSerLysGluLysLeuAspAlaLeuValAspLysCysLysAspArgTyrSerLeu 306
Db      541 TTCTGGCTAGAGAGAGCGAGCAACTCATGAGAAAGAGAGCATTCCTTCCCTC 600
QY      307 ThrTyrMetAlaValAsnLysAspGlySerTyrLysSerAsnValGlyInThrAspVal 326
Db      601 ACATACATCCCGGTGAACAGAGAGGATCATGATCAAAACATCCCTGGAGACGCCG 660
QY      327 AsnAlaValThrTyrGlyValPhe-ProAlaLysGluLleIleGln-ProThrLleValA 346
Db      661 AACCTTCACTTGGGGCGGTCCCGGCAAGGAGATCATCCAGCCCTACCGTGTG 720
QY      346 sPrProValSerPheAsnValTyr-LysAspGluAlaPheGluLleTyrSerArgLysTyr 365
Db      721 ACTGGCGAGCTTATGCTTGGAAAAAGATGAACGTTTACATCTGCTCAGGAGATGG 780
QY      366 Ala-SerLysTyrProGluAspGluAlaSerArgLysLeuValGluGluValGlyLys 385
Db      781 GCCCTGCTGTCCCGAAGGATCATCCAGGAGATTTCTTAACCAACAGTCGCAAAAA 840
QY      385 rHisPheLeuValSerLeuValAspAsnAspTyrLleAsnGlyAspLeuPhe 402
Db      841 CTTATGCTGTGCTGAGCTTGG-GACAAATGACTACCTTAACGGGGACCTTTT 891

```

RESULT 5  
 BG418616  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS

BG418616 823 bp mRNA linear EST 23-OCT-2001  
 HVSMEK0023106f Hordeum vulgare testa/pericarp EST library  
 HVCDNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0023106f, mRNA  
 sequence.  
 BG418616  
 EST  
 Hordeum vulgare.  
 Hordeum vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 ; Triticeae; Hordeum.  
 1 (bases 1 to 823)  
 Wang, R., Close, T. J., Kleinhofs, A., Wise, R., Kannangara, G., von  
 Weistein, D., Akhunov, E., Chin, A., Choi, D. W., Fenton, R. D., Kianian  
 , P., Otto, C., Simons, K., Zhang, D., Begum, D., Frisch, D., Yu, Y.,

TITLE  
 JOURNAL  
 COMMENT  
 FEATURES  
 source

Henry, D., Palmer, M., Rambo, T., Simmons, J., Oates, R. and Main, D.  
 Development of a genetically and physically anchored EST resource  
 for barley genomics: Morex testa/pericarp cDNA library  
 Unpublished (2001)  
 Contact: Wing RA  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Total bp bases = 590  
 Seq primer: AATTAAACCTCACCAAGG  
 High quality sequence stop: 759.  
 Location/Qualifiers  
 1..823  
 /organism="Hordeum vulgare"  
 /cultivar="Morex"  
 /db\_xref="taxon:4513"  
 /clone\_lib="HVSMEK0023106f"  
 /HVCDNA0013 (normal)  
 /tissue\_type="testa/pericarp"  
 /lab\_host="TTC121"  
 /note="Vector: lambdaZAP; Site-1: EcoRI; Site-2: XhoI;  
 plants were raised from seeds in a controlled environments  
 growth chamber maintained in continuous light at 18°C, and  
 testa and pericarp were dissected from developing kernels  
 at Washington State University, Pullman, WA (Kannangara,  
 von Weistein). Total RNA was prepared, poly(A) RNA was  
 purified, one cDNA library was made, and 1 million plu  
 were in vivo excised to give plasmid SK(-) cDNA  
 phagemids in the T7 Close lab at the University of  
 California, Riverside (Akhunov, Chin, Choi, Close, Fenton,  
 Kianian, Otto, Simons, Zhang), phagemids were plated and  
 picked at the Clemson University Genomics Institute (CUGI)  
 (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA  
 preparations, DNA sequencing and sequence analysis were  
 performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates  
 , Rambo, Main). The sequence has been trimmed to remove  
 vector sequence and contains a minimum of 100 bases of  
 phred value 20 or above. For more details on library  
 preparation and sequence analysis see  
 http://www.genome.clemson.edu/projects/barley. To order  
 this clone see http://www.genome.clemson.edu/orders Also  
 see Close TJ, Wing R, Kleinhofs A, Wise R (2001)  
 Genetically and physically anchored EST resources for  
 barley genomics. Barley Genetics Newsletter 31:29-30.  
 (http://wheat.pw.usda.gov/g9pages/bgn/31/cover.html)"

BASE COUNT 236 a 187 c 207 g 192 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4.59e-127 Length: 823  
 Score: 1128.00 Matches: 213  
 Percent Similarity: 85.31% Conservative: 23  
 Best Local Similarity: 77.17% Mismatches: 38  
 Query Match: 51.58% Indels: 2  
 DB: 12 Gaps: 1

US-09-720-451-6 (1-408) x BG418616 (1-823)

QY 68 ASPASNGLUGLUAlaValLysAlaTyrGlyLleHsLeuGlyThrGluMetCysLysLys 87  
 Db 2 GACAATGAGAGAGCTGTGAAGACATATGGAATCCACTTGTACTGAGTGTGCAAGAA 61  
 QY 88 lleLeuAlaHisGlyLleLysThrLeuHisLeuTyrThrLeuAsnMetGluLysSerAla 107  
 Db 62 AATTTGGCTAGTGGGATCAAGACTTGCACCTTACACTTAACATGCAAGAACCGCT 121  
 QY 108 LeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSerLysValSerArgSerLeuPro 127  
 Db 122 TTAGCAATCTGATGATCTTGATTAATAGAGGAATCCAGAGCTTCAAGACATTAACCT 181



```

Db      184 CTTGAGCCAGAGCAAGATAATGAGCTGCTGAAAAGATTACACAAAGGCGCTTC 243
QY      251 LeuThrIleAsnSerGlnProAlaValAsnGlyIuLysSerAspSerProThrValGly 270
Db      244 CTCACCAATCAATACCCAGCCAGCTGTCAATGGGAAAAGTCAGATCTCTCTACTGTTGGC 303
QY      271 TTPGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
Db      304 TGGGTGGACACAGGGGGTATGTTTACCAGAACCATGTAGAGATTCTCTGCTCAAG 363
QY      291 GluLysIleuAspAlaLeuValAspLysCysLysAspArgTyrSerLeuThrTyrMetAla 310
Db      364 GAAAAGCTGAGATGCACCTTGTGATTAATGACAAGATCCAAACATCTTTAACTTAATAGGCT 423
QY      311 ValAsnLysAspGlySerTyrLysSerAsnValGlyGlnTrrAspValAsnAlaValThr 330
Db      424 GTGAATTAAGACGGAGAGCTGGAAATCTAATGTGGGTCAAACTGATGATGCTGTGACA 483
QY      331 TTPGlyValAlaPheProAlaLysGluIleIleGlnProThrIleValAspProValSerPhe 350
Db      484 TGGNGTGTCTTCCAGCTAAGAGATTAATCAACCAACCATTTGATCTGTGACGCTTC 543
QY      351 AsnValTrrPylsAspGluAlaPheGluIleTrrPserArgGlyTrrPalaSerLeuThrPro 370
Db      544 AATGATGGAAGATGAGNCATTTGATAAATTTGGTCAAGAGATGGGCGACGTTGCCCT 603
QY      371 GluAspGluAlaSerArgLysLeuValGluGluValGlyGlySerHisPheLeuValSer 390
Db      604 GAGGGTGATGCATTCACGAAATTTGGTGAAGAATTTGGGGGACGCTTACTTCTGTGAGT 663
QY      391 Leu 391
Db      664 TTG 666

RESULT 7
LOCUS   BQ994876              760 bp      mRNA      linear      EST 21-AUG-2002
DEFINITION   QGF8G21.yg.ab1 QG.EF8HJ lettuce serriola Lactuca sativa cDNA clone
ACCESSION   BQ994876
VERSION     BQ994876.1 GI:22414411
KEYWORDS
SOURCE      Lactuca sativa.
ORGANISM    Lactuca sativa.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE   1 (bases 1 to 760)
            Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
            ,P., Kojkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
            lettuce and sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
            unpublished (2002)
JOURNAL     Contact: Alexander Kozik [R.W.Michelmore]
COMMENT     Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@catc.ucdavis.edu [michelmore@vegmail.ucdavis.edu]
            belongs to contig QG_CA.Contig1724, see http://cgpdb.ucdavis.edu/
            for details.
FEATURES
            Source          location/Qualifiers
            1..760
            /organism="Lactuca sativa"
            /cultiivar="L.serriola"
            /db_xref="taxon:4236"
            /clone="QGf8G21"

```

```

/clone.lib="QG.EF8HJ lettuce serriola"
/lab_host="E.coli"
/Note="Vector: pBRCDNASFIAB. The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB=QG.EF8HJ lettuce serriola
TAG_TISSUE=flowers post-fertilized
TAG_SEQ=GCCATCGGG"

BASE COUNT      226 a      167 c      175 g      192 t
ORIGIN

Alignment Scores:
Pred. No.:      2,516-123      Length:      760
Score:          1097.00      Matches:      198
Percent Similarity: 88.89%      Conservative: 26
Best Local Similarity: 78.57%      Mismatches:  28
Query Match:    50.16%      Indels:      0
DB:              14      Gaps:      0

US-09-720-451-6 (1-408) x BQ994876 (1-760)
QY      35 IleMetProIleAsnAsnTyrLysGlyPheIleAlaArgMetThrGlyPheCysLysThrLys 54
Db      3 ATCATGGCCCAATTAACAACCTTAATGATGTTCAATGAGTAGCTGTTCTTCCAAAATAAG 62
QY      55 IleProAlaAspIleMetAlaAlaLeuGluProIleLysAspAsnGluAlaValLys 74
Db      63 ATTCATCTGTGAATTAATTAATGCTGCTTGGAGCCTATCAAGACCAATGAGAGAGCTGTAGA 122
QY      75 AlaTyrGlyIleHisLeuGlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLys 94
Db      123 AACATGAGATGCCACCTTGGAACTGAAATGTGTAGAAGAAATCAAGGAGAGTGAATCAAG 182
QY      95 ThrLeuHisLeuTyrThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeu 114
Db      183 ACTTGTGATCTTTTATACCTCAACACATGAGAAATCTGCATTGGCAATTTGACAAAGTCT 242
QY      115 GlyLeuIleGluGluSerLysValSerArgSerLeuProTrrArgArgProAlaAspVal 134
Db      243 GGATTTGGTTGACGAGGCCAAATTAACAAGACCACCTTCATGAGAGAGACCTTAACCTT 302
QY      135 PheArgValLysGluAspValArgProIlePheThrPalaAsnArgProLysSerTyrIle 154
Db      303 TTCCGCTTAAAGAGATGTTAGGCCAATTTTCTGGGCCAATCGTCCAAAGAGCTTAACATA 362
QY      155 SerArgThrIleGlyTrrAspGlnTyrProHisGlyArgTrrPglYAspSerCysAsnPro 174
Db      363 TCAAGGAGCGGTGGTGGAGCAATATCACATGGCGCTTGGGGGATTTCTTAATGCA 422
QY      175 SerTyrGlyAlaLeuSerAspTyrGlnPheMetArgProArgAlaArgAspLysLeu 194
Db      423 TCATACGGAGCATTTAACGATCATCATGAGCCAGCGACGTCGAGCAAGAAATTT 482
QY      195 ValGluGlnTrrPalaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeu 214
Db      483 CAAGCAAGATGGGTGCTTGTGACCAAGTTTGATGATTAATTTCCGAGAAATTCATGAAG 542
QY      215 TyrCysLeuGlyLysLeuArgSerAsnProTrrSerGlnLeuAspGlyLeuGlnProGlu 234
Db      543 TACTGCTTGGGCACTAATAAACACAGCCCTGTGTGGACCTGATGAGCGCTTCAGCCAAG 602
QY      235 ThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsn 254
Db      603 ACAAAATCATCAACGAGAGACTGGCTTCCATCAACCTCAAGGTTTCTCAACATCAAC 662
QY      255 SerGlnProAlaValAsnGlyLysSerAspSerProThrValGlyTrrPylsGlyPro 274

```

Db	663	AGAAACCCGCTGTATATGAGCAAAATCTGACATCCCAATCTGTGATGGGCCAGCG	722
Oy	275	GLYCYTYRValTyrGlnLysAlaTyrValGlnPhe	286
Db	723	GGTGGATACCTATATACGAAAGCCTACGTGAGATT	758
RESULT 8			
LOCUS	AM830074		
DEFINITION	631 bp MRNA linear EST 03-DEC-2001		
ACCESSION	SM22102.Y1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:		
VERSION	Gm-cl028-4684.5' similar to SM:U761.ARA7H.080585 PROBABLE		
KEYWORDS	METHYLENETHETRAHYDROFLATE REDUCTASE ;, mRNA sequence.		
SOURCE	AM830074		
ORGANISM	AM830074.1 GI:79240066		
REFERENCE			
AUTHORS	Glycine. 1 (bases 1 to 631)		
TITLE	Shoemaker, R., Kelm, P., Vodko, L., Erpelding, J., Corryell, V., Khanna		
JOURNAL	, A., Bolla, B., Marraz, M., Hillier, L., Kucaba, T., Matlin, J., Beck, C.,		
COMMENT	Wille, T., Underwood, K., Stepec, M., Theising, B., Allen, M., Bowers		
	, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck		
	, R., Ralter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann		
	, R., Waterston, R. and Wilson, R.		
	Public Soybean EST Project		
	Unpublished (1999)		
	Contact: Shoemaker R./Public Soybean EST Project		
	Public Soybean EST Project		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: estewatson.wustl.edu		
	This clone is available through: Resgen, Invitrogen Corp, 2130		
	South Memorial Parkway Huntsville, AL 35801 For further information		
	call: (800)-533-4363 or contact via email: coueresgen.com		
	Insert Length: 1257 Std Error: 0.00		
	High quality sequence stop: 447.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..631		
	/organism="Glycine max"		
	/db_xref="taxon:3847"		
	/clone="GENOME SYSTEMS CLONE ID: Gm-cl028-4684"		
	/clone_11b="Gm-cl028"		
	/tissue-type="roots of 'Supernod' plants"		
	/lab_host="DH10B"		
	/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:		
	XhoI; The mRNA was isolated from roots of Glycine max		
	'Supernod' plants generously donated by Dr. Gary Stacey.		
	The seedlings were inoculated with Bradyrhizobium		
	japonicus, strain USDA110 prior to harvest. Striagene's		
	Japonicus synthesis kit (catalog number 200401) was used to		
	synthesize the cDNA. First-strand synthesis was performed		
	with 5-methyl dCTP, hence the ligated cDNA was		
	hemimethylated. A modification of Striagene's		
	first-strand synthesis primer was used. An 'anchor'		
	nucleotide (V-A-C, or G) was added to the 3' end of the		
	primer [GAGAGAGAGAGAGAGAGAGACTGATCTCGAG(T)18V] to anchor		
	the primer at the 5' end of the poly(A) tract. After		
	second-strand synthesis, the cDNA ends were filled in with		
	cloned Pfu DNA polymerase, ligated to EcoRI adaptors and		
	subsequently phosphorylated. The XhoI site within the		
	first-strand synthesis primer was then restricted by		
	digestion with XhoI; all XhoI sites in the cDNA would be		
	protected by their hemimethylated status. The cDNA		
	constructs were size-fractionated with a 500bp cutoff.		
	using GblcooBRI life technologies' cDNA Size Fractionation		
	column. The column eluent was then ligated into		
	Striagene's pBluescript II XR Predigested vector		

BASE COUNT	191 a	114 c	158 g	166 t	2 others					
ORIGIN	Coryell. <sup>a</sup>									
(pBluescript II SK(+) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Kelm and Dr. Virginia										
Alignment Scores:										
Pred. No.:	5	65e-123	Length:	631						
Score:	1093.00	Matches:	202							
Percent Similarity:	96.67%	Conservative:	1							
Best Local Similarity:	96.19%	Mismatches:	7							
Query Match:	49.98%	Indels:	0							
DB:	10	Gaps:	0							
US-09-720-451-6 (1-408) x AM830074 (1-631)										
QY	138	LysLlnuaspAlaArgProIlePheThrAlaAsnArgProLysSerYrIleSerArgThr	157							
Db	2	AAAGAAAGATCCGCTCCAACTTTTGGCAAACTCCTCCAAAACATCAATACAGAC	61							
QY	158	IleGlyTrpaspGlnTrpProHisGlyArgTrpGlyAspSerCysAsnProSerYrGly	177							
Db	62	ATAGAGATGGATCAATACACACAGCGCGGTGGGGTGATCTGTATCCATCATATGCT	121							
QY	178	AlaIeuSerAspYrGlnPheMetArgProAlaArgAspLysLysLeuValGluGlu	197							
Db	122	GCATTATCTATTTATTCAGTTTCGCGCACGCTGCACGGGACAGACAGCTTTGTAAGAA	181							
QY	198	TrpAlaValProLeuLysSerValGluAspIleYrGluArgPheArgLeuYrCysLeu	217							
Db	182	TGGGCACTTCCCTGGAAAAGCATGTGAAGTATATGTGAGAGCTTTAGACTGATGTGCTT	241							
QY	218	GlyLysLeuArgSerAsnProTrpSerGluLeuAspGlyLeuGlnProGluThrLysIle	237							
Db	242	GGAAAGTTGAGAACCATCTTGTCAGACACTAGTGGTCTTCAGCCAGACAGAAAGATA	301							
QY	238	IleAsnGlnLeuGlnLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnPro	257							
Db	302	ATAATAGCTGCTGGAAAAGATTAACACAAAGGGCTCCCTCACCAATCAATACACGCA	361							
QY	258	AlaValAsnGlyLysSerAspSerProThrValGlyTrpGlyGlyProGlyGlyYr	277							
Db	362	GCCTCAATGGGAAAGATCATTTCTCTCTCTGTTGGCTGGGCTGGACAGCGGGTAT	421							
QY	278	ValTrpGlnLysAlaYrValGluPhePecySerLysGluLysLeuAspAlaLeuVal	297							
Db	422	GTTTACCAAGGACAAATATGAGAGTCTTCTCTTAAGGAAAAGCTGGATGACACTGTGT	481							
QY	298	AspLysCysLysAspArgThrSerLeuThrYrMetAlaValAsnLysAspGlySerTrp	317							
Db	482	GATTAATGCAAGGATCGAACATCTTAACCTATATAGGCTGTGATTAAGACGGGACCTG	541							
QY	318	LysSerAsnValGlnThrAspValAsnAlaValThrTrpGlyValPheProAlaLys	337							
Db	542	AAATCAATATGGGCTAAACTGATGTGTAATGCTGTGACATGGGGTGTCTTCCAGCTTAG	601							
QY	338	GluIleIleGlnProThrIleValAspPro	347							
Db	602	GAGATTATTCAAACCAACATTTGTGATCCT	631							
RESULT 9										
LOCUS	BE576981	700 bp	MRNA	linear	EST 20-FEB-2001					
DEFINITION	NCI-1588r3 Ice plant lambda Uni-2ap XR expression library, 0 hours									
FEATURES	mRNA sequence.									
ACCESSION	BE576981									
VERSION	EST.	GI:9826689								
KEYWORDS	common iceplant.									

ORGANISM Mesembryanthemum crystallinum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllidae; Caryophyllales; Alzooceae; Mesembryanthemum.  
REFERENCE 1 (bases 1 to 700)  
AUTHORS Cushman, J. C.  
TITLE An expressed sequence tag database for the common ice plant,  
Mesembryanthemum crystallinum  
JOURNAL Unpublished (1997)  
COMMENT Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu  
PCR Primers  
FORWARD: T7  
BACKWARD: T3  
Plate: L0-16 row: H column: 4  
Seq primer: T3  
High quality sequence stop: 350  
POLYA-No.

FEATURES  
source location/Qualifiers  
1..700  
/organism="Mesembryanthemum crystallinum"  
/db\_xref="taxon:3544"  
/clone="L0-1588"  
/clone\_lib="Ice plant Lambda Uni-zap XR expression library  
, 0 hours NaCl treatment"  
/tissue\_type="leaf"  
/dev\_stage="Six week old"  
/note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI"  
BASE COUNT 226 a 136 c 165 g 173 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.66e-122 Length: 700  
Score: 1087.00 Matches: 198  
Percent Similarity: 92.70% Conservative: 18  
Best Local Similarity: 84.98% Mismatches: 17  
Query Match: 49.70% Indels: 0  
DB: 10 Gaps: 0

US-09-720-451-6 (1-408) x BB576981 (1-700)

QY 48 ThrGlyPheCysLysThrLysIleProAlaSerIleMetAlaIleLeuGluProIleLys 67  
|||||  
DB 1 ACTGTTTCTGTAAACAAGATTCTCAAGAGTTACTGATGCATTGGAACCAATTAG 60  
|||||  
QY 68 AspaSngLgUAlaValLysAlaTyrGlyIleHisLeuGlyThrGluMetCysLysLys 87  
|||||  
DB 61 GACAAATGATGAGCGCGTGAAGCCCTACCGGATTCACCTCGGAACATGGAATGTGCAAGAG 120  
|||||  
QY 88 IleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThrLeuAsnMetGluLysSerAla 107  
|||||  
DB 121 ATTGTGGCTAGTGAATTAACAAGTTGCAATGATATACCTCGAATATGACAAACTCA 180  
|||||  
QY 108 LeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSerLysValSerSerLeuPro 127  
|||||  
DB 181 TTAGCAATTTTGAAGATCTTGATTAATTGAAGATCCAAATCCAAAGTCATTACT 240  
|||||  
QY 128 TTPATGATGProAlaAsnValPheArgValLysGluAspValArgProIlePheTrpAla 147  
|||||  
DB 241 TGGAAACGCCCTGCAAAATGTGAACCGTTAAAGAAAGATGTCGCCATATTTGGGCA 300  
|||||

QY 148 AsnArgProLysSerTyrIleSerArgThrIleGlyTTPAspGlyTyrProHisGlyArg 167  
|||||  
DB 301 AATTCGTCGCAAAAGTTACTTACGGGACAGTTGATGGATGACATACCCACAGGGGGA 360  
|||||  
QY 168 TTPGlyAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyrGlnPheMetArgPro 187  
|||||

DB 361 TGGGATGATTCGCAATGCTTCTTATGAGGACATTAGTACCATTATGAGGCCA 420  
QY 188 ArgAlaArgAspLysLysLeuValGluGluTTPAlaValProLeuLysSerValGluAsp 207  
|||||  
DB 421 CGTGGCCGTGAATGAAGGCTCCAAAGAGAGTGCGTTGCCACTAAAGTGTGAAGAT 480  
QY 208 IleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeuAsnSerAsnProTyrSerGlu 227  
|||||  
DB 481 ATACATGAGAGGTTCCAAACTACTCTCTGTTAAGCTAAAGTAATCTGCTGTA 540  
QY 228 LeuAspGlyLeuGluInProGluThrLysIleLeuSngLgUleuGluLysIleAsnThr 247  
|||||  
DB 541 ATAGATGGCCCTTACGACAGAAACAGATCATCAATGAACAGCTGGCCAAATATAG 600  
QY 248 LysGlyPheLeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspSerPro 267  
|||||  
DB 601 AAGGGCTTCTCACCATCAACACACCACTGCTGTTAATGAGAAATCTACTCCCA 660  
QY 268 ThrValGlyTTPGlyLysGlyProGlyGlyTyrValTyrGln 280  
DB 661 AGTGTGATGGGAGAGCTCTGATGATATGTAATACCA 699

RESULT 10  
BI421654  
LOCUS EST532320 tomato callus, TAMU Lycopersicon esculentum cDNA clone  
DEFINITION CLEC67D19 5' end, mRNA sequence.  
ACCESSION BI421654  
VERSION BI421654.1 GI:15195373  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE 1 (bases 1 to 752)  
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,  
'Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Roming,  
'C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
Generation of ESTs from tomato callus tissue  
Unpublished (1999)  
JOURNAL  
CONTACT: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.  
Location/Qualifiers  
1..752  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEC67D19"  
/clone\_lib="tomato callus, TAMU"  
/tissue\_type="callus"  
/dev\_stage="25-40 days old"  
/lab\_host="XLI-Blue MRF"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; supplier: Giovannoni laboratory; cDNA - Cotyledons  
of seedlings 7-10 days post-germination were excised, cut  
at both ends and placed on MS medium with no selection.  
Mixed callus was harvested at 25 and 40 days and included  
undifferentiated masses. Tomato Callus EST Library"

FEATURES  
source  
1..752  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEC67D19"  
/clone\_lib="tomato callus, TAMU"  
/tissue\_type="callus"  
/dev\_stage="25-40 days old"  
/lab\_host="XLI-Blue MRF"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; supplier: Giovannoni laboratory; cDNA - Cotyledons  
of seedlings 7-10 days post-germination were excised, cut  
at both ends and placed on MS medium with no selection.  
Mixed callus was harvested at 25 and 40 days and included  
undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 216 a 164 c 185 g 187 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.42e-121 Length: 752  
Score: 1079.50 Matches: 204  
Percent Similarity: 89.24% Conservative: 20  
Best Local Similarity: 81.27% Mismatches: 27  
Query Match: 49.36% Indels: 1  
DB: 13 Gaps: 0

```

US-09-720-451-6 (1-408) x B418214 (1-752)
OY      88  ILEUVALAHISGLYILEYSTRHLEUHSLEUTYRTHRLAUSMECTLYLSERALA 107
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      2  ATTTTAGCCATGCGCTGCAAGACCTTGCTTTTATCATTAACATGAGAAATCACCA 61
OY      108  LEUVALIIELEUMETASNDLEUGLYLEULIEGLUGLUSERLYSVALSERARSERLEUPRO 127
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      62  TTGGCCATTTTGTGATCTTGATTAATCGAAGGTCCTCAAAATTTCTAGGCCATTGGCT 121
OY      128  TTPATGATGPROVALASNDVALPHEARGVALLYSGUASPVALARGPROLLEHETTPALA 147
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      122  TGGAGACGTCTGCAAAATTTTCCGTGTCAGAGGAGATGCTCCTATATTATTGGGCG 181
OY      148  ASNATGPROLYSSEYTRILLESERATGTHRIEGLYTPASPGLNTHPROHISGLYARG 167
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      182  AATCGCCAAAGACGTACTTTCANAGACACAGGTTGGGATGAATACCCACATGTGAGA 241
OY      168  TTPGLYASPSEYSCYSANPROSEYTRGLYALALEUSERPYPYGLNPHEMETARPRO 187
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      242  TGGAGCAATGCTCAAAATTCATCATACGAGCATTTACCGACTATCAGTTCATGCGTGA 301
OY      188  ARGALATARGASPULYSLEUVALIGLUGLUTTPALAVALPROLEULYSERYALGLUASP 207
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      302  CGCTCAAGATGATAGAGCTTCAGAGAGATGGCTGCGGCTTGAAATAGCGTGGAGAT 361
OY      208  ILETYGLUARGPHEARGLEUTYRCYLSLEUGLYLYLSLEUARGSERANPROTRPSERGLU 227
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      362  ATCTGACGAAATTCAGAGACTACTGTCTTGAAAGCTGAGAGAGTGGCCATGCTCTGAG 421
OY      228  LEUASPGLYLEUGLNPROGLUTHRILYSILELEASNGLUGLNULEUGLYLSILEASNTHR 247
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      422  TTAGATGGCTTCCACCGACGACAAAGATCATAAACGAAAGTGGTCAATGCTCAACAG 481
OY      248  LYSGLYPHELEUTHRIEASNSERGLNPROVALASNGLYGLULYSERPASPSEYRPO 267
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      482  AAGGCTTCTCTACTATTTAAACGACCAACGAGATTAATGCTGAGAAATCCGACTCTCT 541
OY      268  THRVALIGLYTPGLYGLYPROGLYGLYTYRVALTYRGLNLYSALATYRVALGNPHEPHE 287
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      542  TCTGTTGGATGGGTGGCCCGCGCGGATATGTTTATCAAAAGGCTACTTGAGATCTCTTC 601
OY      288  CYSSEYRGLULYSLEUASPALALEUVALASPLYSCYLSYASPAVGYTHSERLEUTHR 307
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      602  TGTCTTCAGGAGAAATTTAAACGCCCTCTGTGAGAAATCAAGTCTTCCCTTACCCCTCA 661
OY      308  TYRMETALAVALASNDLYSASPGLYSEYTRPYSSERASNDVALIGLINTHRASPVALASN 327
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      662  TACATGGGCTGTACACGAGAGGAACTGGGATCTTCCACGTCAC--CCAAACCGACATCAAC 719
OY      328  ALAVALTHTTTPGLYVALPHEPROVALALYSGLU 338
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      720  GCTGTGACATGGGAGTTTCCACGACTAAGAA 752

RESULT 11
LOCUS   B418214 842 bp mRNA linear EST 23-OCT-2001
DEFINITION HVSMEK0021017f Hordeum vulgare testa/pericarp EST library
            HVCNMA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0021017f, mRNA
            sequence.
ACCESSION B418214
VERSION   B418214.1 GI:13323765
KEYWORDS  EST.
SOURCE    Hordeum vulgare.
ORGANISM  Hordeum vulgare.
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
          ; Triticeae; Hordeum.
          1 (bases 1 to 842)
REFERENCE  Ming R., Close T.J., Kleinof A., Wise R., Kannangara G., von
          Weinstein, D., Akhunov E., Choi D.W., Fenton R.D., Kianian
          P., Otto C., Simons R., Zhang D., Begum D., Frisch D., Yu Y.,

```

```

TITLE      Henry D., Palmer M., Rambo T., Simmons J., Oates R. and Main D.
            Development of a genetically and physically anchored EST resource
            for barley genomics: Morex testa/pericarp cDNA library
JOURNAL    Unpublished (2001)
COMMENT    Contact: Ming RA
            Clemson University Genomics Institute
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: twing@clemson.edu
            Total hg bases = 539
            Seq primer: AATTAACCTCACTAAGG
            High quality sequence stop: 641.
FEATURES   location/Qualifiers
            source
            1..842
            /organism="Hordeum vulgare"
            /cultivar="Morex"
            /db_xref="taxon:4513"
            /clone="HVSMEK0021017f"
            /clone_1lb="Hordeum vulgare testa/pericarp EST library
            HVCNMA0013 (normal)"
            /tissue-type="testa/pericarp"
            /lab_host="TUC121"
            /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
            Plants were raised from seeds in a controlled environments
            growth chamber maintained in continuous light at 18°C, and
            testa and pericarp were dissected from developing kernels
            at Washington State University, Pullman, WA (Kannangara,
            von Weinstein). Total RNA was prepared, poly(A) RNA was
            purified, one cDNA library was made, and 1 million pfu
            were in vivo excised to give plasmid SK(-) cDNA
            phagmids in the rt close lab at the University of
            California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
            Kianian, Otto, Simons, Zhang). Phagmids were plated and
            picked at the Clemson University Genomics Institute (CUGI)
            (Begum, Palmer, Frisch, Atkins and Ming). Plasmid DNA
            preparations, DNA sequencing and sequence analysis were
            performed at CUGI (Ming, Yu, Frisch, Henry, Simmons, Oates
            , Rambo, Main). The sequence has been trimmed to remove
            vector sequence and contains a minimum of 100 bases of
            phred value 20 or above. For more details on library
            preparation and sequence analysis see
            http://www.genome.clemson.edu/projects/barley. To order
            this clone see http://www.genome.clemson.edu/orders Also
            see close rt, Ming R., Kleinof A., Wise R (2001)
            Genetically and physically anchored EST resources for
            barley genomics. Barley Genetics Newsletter 31:29-30.
            (http://wheat.pw.usda.gov/ggpages/dgn/31/cover.html)"
BASE COUNT 238 a 193 c 215 g 195 t 1 others
ORIGIN
Alignment Scores:
pred. No.:      8,39e-118      Length:      842
Score:          1052.50        Matches:      209
Percent Similarity: 84.25%      Conservative: 21
Best Local Similarity: 76.56%      Mismatches: 42
Query Match:    48.13%        Indels:      3
DB:             12            Gaps:        1

US-09-720-451-6 (1-408) x B418214 (1-842)
OY      70  GLUGLVALAVALSALATYRGLYILEHSLEUGLYTHRGLUMETCYLSYLSLEULEU 89
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      2  GAGGAGGCTGTGAAACCATATGGAATCCACTTGGACTGAGATGGCAAGAAATTTTGG 61
OY      90  ALAHISGLYILEYSTRHLEUHSLEUTYRTHRLAUSMECTLYLSERALALEUALA 109
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      62  GCTAGTGGGATCAAGACTTTGACACCTTTACACACTAAACATGAGAAAGCCCTTAGCA 121
OY      110  ILEUMETASNDLEUGLYLEGLUGLUSERLYSVALSERARSERLEUPROTTPARG 129
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB      122  ATTTGATGAATCTTGATTAATAGAGAAATCCAAAGCTTTCAAGAAACATTACTTGAGAG 181

```

QY 130 ArgProAlaAsnValPheArgValIleGluAspValArgProIlePheTrpAlaAsnArg 149  
 DB 182 CCACCACTAATGTTTCCGTCGTAAGAGATGTCGCCCTATATCTGGCCACACAGA 241  
 QY 150 ProLysSerTrpIleSerArgThrIleGlyTrpAspGlnTrpProHisGlyArgTrpGly 169  
 DB 242 CCAAGAGTACATTCAAGAGCAGCTGGGATCAATACACAGAGGGGGGCTGGGGT 301  
 QY 170 AspSerCysAsnProSerTrpGlyValAlaLeuSerAspTrpGlnPheMetArgProArgAla 189  
 DB 302 GATTTCAGAGAACCTCATACAGAGCCCTTATATACACAGCTTCAACAGCCGACCTGGG 361  
 QY 190 ArgAspLysIleValGluGlnTrpAlaValProLeuLysSerValGluAspLysTrp 209  
 DB 362 CATTGTAAGAAGCTCCAGAGGAATGGGCTGTCACAGAAATCTGTGCAAGACATTAAT 421  
 QY 210 GluArgPheArgLeuTrpCysLeuGlyLysLeuArgSerAsnProTrpSerGluLeuAsp 229  
 DB 422 GAGCGGTTCTGTAACCTCTGTCAAGAAACTTAAAGCACCCATGGCTGAGATTAGAT 481  
 QY 230 GlyLeuGlnProGluThrLysIleIleAsnGlnLeuGlnLysIleAsnThrLysGly 249  
 DB 482 GGCTTCAACCGCAGACAAAGATATGATGATGATGATGATGATGATGATGATGATGAT 541  
 QY 250 PheLeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrVal 269  
 DB 542 TTCTTACATCATACAGCCACCTGCTGTAATGCAAGAAATGATGATGATGATGATGAT 601  
 QY 270 GlyTrpGlyValProGlyLysTrpValTrpGlnLysAlaTrpValGluPheCysSer 289  
 DB 602 GGATGGGGGGCTCCAGAGAGCTATGTTTACCA -AAGGCTACGTCGATTTCTTGGCGCT 660  
 QY 290 LysGluLysLeuAspAlaLeuValAspLysCysLysAspArgThrSerLeuThrTyrMet 309  
 DB 661 AAGGACAGAGTGGACCACTCATCGAAGAAAGAGAGGCGATTCCTTCCATCATACATC 720  
 QY 310 AlaValAsnLysAspLysSerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329  
 DB 721 GCGGGGAGCAAGAGGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 777  
 QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnPro 342  
 DB 778 CTTTGGGGCGTGGTCCCGG -CAAGAAATCATCCACCT 815  
 RESULT 12  
 BE805904 611 bp mRNA linear EST 06-DEC-2001  
 LOCUS ss62507.y1 Gm-c1062 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-c1062-1046 5' similar to SW:MTHR\_ARATH O80585 PROBABLE  
 METHYLENETETRAHYDROFOLATE REDUCTASE ; mRNA sequence.  
 BE805904  
 ACCESSION BE805904.1 GI:10237016  
 VERSION  
 KEYWORDS  
 SOURCE soybean.  
 ORGANISM  
 EST.  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 611)  
 REFERENCE  
 AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna  
 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
 R., Waterston,R. and Wilson,R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cou@resgen.com  
 Insert length: 1101 Std Error: 0.00  
 High quality sequence stop: 410.  
 Location/Qualifiers  
 1. 611  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1062-1046"  
 /tissue\_type="stem tissue of greenhouse grown plants"  
 /dev\_stage="1 month old"  
 /lab\_host="DH10B"  
 /note="Vector: Bluescript II SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from stem tissue of 1 month old greenhouse grown plants  
 for the cultivar Raiden. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with a XhoI restriction site. EcoRI  
 adapters were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA fragments were  
 directionally cloned into the EcoRI-XhoI restriction site  
 of the Bluescript vector. The ligated cDNA fragments were  
 transformed into DH10B host cells (GibcoBRL). This library  
 was constructed in the laboratory of Dr. Randy Shoemaker."  
 BASE COUNT 178 a 101 c 165 g 166 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,61e-118 Length: 611  
 Score: 1052.00 Matches: 196  
 Percent Similarity: 98.03% Conservative: 3  
 Best Local Similarity: 96.35% Mismatches: 4  
 Query Match: 48.10% Gaps: 0  
 DB: 12  
 US-09-720-451-6 (1-408) x BE805904 (1-611)  
 QY 202 LeuLysSerValGluAspIleTrpGluArgPheArgLeuTrpCysLeuGlyLeuAsnArg 221  
 DB 3 TTGAAAAGCATTCAGATATATATAGAGAGTTTAACTGTATGTCCTTGAAGATTGAGA 62  
 QY 222 SerAsnProTrpSerGluLeuAspGlyLeuGlnProGluThrLysIleIleAsnGln 241  
 DB 63 AGCAATCCTTGTCACAGACTAGTGTCTTCAAGCAGACAAAGATTAATGAGCTG 122  
 QY 242 LeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsnGly 261  
 DB 123 CTGGAAGAAAGATTAAACAAAGGGCTTCCTCCATCAATAGGCACAGCTGTCAATGG 182  
 QY 262 GluLysSerAspSerProThrValGlyTrpGlyGlyProGlyLysTrpValTrpGlnLys 281  
 DB 183 GAAAGTCAATTCCTCTAATGTTGGCTGGGCGGACCAAGCGGGATGTTACAGAAAG 242  
 QY 282 AlaTrpValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCysLys 301  
 DB 243 GCATATGATAGAGTCTTCTCTCTAAGAAAGCTGAGTGCATGTTATTAATGCAAG 302  
 QY 302 AspArgThrSerLeuThrTrpMetAlaValAsnLysAspGlySerTrpLysSerAsnVal 321  
 DB 303 GATGAAACATCTTAATCTTAATAGCTGTGATTAAGACGGAGAGCTGGAATCTATGTG 362  
 QY 322 GlyGlnThrAspValAlaAsnAlaValThrTrpGlyValPheProAlaLysGluIleGln 341  
 DB 363 GGTCAAACTGATGTTATGCTGACATGGGCTGCTCCAGCTTAAGAGATTAATTCAA 422  
 QY 342 ProThrIleValAspProValSerPheAsnValTrpLysAspGlnAlaPheGluIleTrp 361  
 DB 423 CCAACCATTTGATCTCTGTCAGCTTCATATGTAAGAGATGAGCATTTGAAATTTGG 482

Qy 362 SerArgGlyTrpAlaSerLeuTyrProGluuAspGluAlaSerArgLysLeuValGluGln 381  
|||||  
Db 483 TCAGAGAGATGGGCAAGCTTGTACCTGAGGTGATGATCCAGAAATTTGGTTGAAGAG 542  
Qy 382 ValGlyGlySerHisPheLeuValSerLeuValAspAsnAspTyrTrileasnGlyAspLeu 401  
|||||  
Db 543 GTTGGGGGACGACTTCTTGGTGTGATGTCGACATGATATCAATCAATGATGATCTT 602  
Qy 402 PheAlaVal 404  
|||||  
Db 603 TTCACCGTC 611  
RESULT 13  
BG123885 729 bp mRNA linear EST 31-JAN-2001  
LOCUS BG123885 tomato shoot/meristem Lycopersicon esculentum cDNA clone  
DEFINITION CTOF3B13 5' sequence, mRNA sequence.  
ACCESSION BG123885 GI:12624073  
VERSION BG123885.1 GI:12624073  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE 1 (bases 1 to 729)  
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Uterback,T.,  
Hansen,C., Roming,C. and Tanksley,S.  
TITLE Generation of ESTs from tomato shoot/meristem tissue  
JOURNAL Unpublished (2001)  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.  
FEATURES  
source  
1..729  
/organism="Lycopersicon esculentum"  
/cultivar="TA96"  
/db\_xref="taxon:4081"  
/clone="CTOF3B13"  
/clone\_11b="tomato shoot/meristem"  
/tissue\_type="shoot/meristem"  
/dev\_stage="developing shoots from 4-6wks old plants"  
/lab\_host="SOLR"  
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:  
XhoI. Small expanding leaves from the growing tip were  
taken from greenhouse plants (4-6wks old TA96). Tissue  
was immediately frozen in liquid nitrogen."  
BASE COUNT 216 a 155 c 180 g 178 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,76e-117 Length: 729  
Score: 1049.00 Matches: 200  
Percent Similarity: 89.75% Conservative: 19  
Best Local Similarity: 81.97% Mismatches: 22  
Query Match: 47.97% Indels: 3  
DB: 12 Gaps: 0  
US-09-720-451-6 (1-408) x BG123885 (1-729)  
Qy 101 LeuAsnMetGlyIuysSerLalaLeuLalaLeuMetAsnLeuGlyLeuGluGlnSer 120  
Db 3 TTAAACATGGAGAAATCAGCATTTGATGAAATCTTGATTAATCGAAGAGATCC 62  
Qy 121 LysValSerArgSerLeuProTrpArgTrpAlaAsnValPheArgValLysGluAsp 140  
|||||  
Db 63 AAAATTTCTAGGCCATTGCCCTTGAGACGCTCTGCAAAATATTTCCGTCGTCAGAGAGAT 122  
Qy 141 ValArgProIlePheTrpAlaAsnArgProLysSerTyrTrileSerArgTrnIleGlyTrp 160

Db 123 GTGGCTCTATATTTGGGGGATGCTCCAAAGACCTACATTTCAAGACCAAGCTTGG 182  
Qy 161 AspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180  
Db 183 GATCAATATCCACATGATGATGATGAGCAATAGCTCAAAATCATCATATACGGACCTTACC 242  
Qy 181 AspTrpGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGlnTrpAlaVal 200  
Db 243 GACTATCATGTTCAATGCGTCGACGCTCAAGAGATTAAGAGCTTCAAGAGATGGCTGCG 302  
Qy 201 ProLeuLysSerValGluAspTyrTrpGlyuArgPheArgLeuTyrCysLeuGlyLysLeu 220  
Db 303 GCTTGAATAGCGTGAAGATATATACGAAATTCAGAACTACTGCTTGGAAACCTG 362  
Qy 221 ArgSerAsnProTrpSerGluLeuAspGlyLeuGlnProGluTrpLysLileAsnGln 240  
Db 363 AGAAGCTGCCCATGCTGCTGATGATGCTTACGCCAGAGACAAAGATCATTAACGAA 422  
Qy 241 GlnLeuGluLysLileAsnTrpLysGlyPheLeuTrnIleAsnSerGlnProAlaValAsn 260  
Db 423 AAGTTGGGTGAGGTCAACAGAAAGGCTTCGACTATTACAGCAACGACGATTAA 482  
Qy 261 GlyIuysSerAspSerProTrpValGlyTrpGlyGlyProGlyGlyTyrValTrpGln 280  
Db 483 GCTGAAGAGTCCGACATCTCTCTGTTGATGGGTGGCCCGCGGATATGTTATCAA 542  
Qy 281 LysAlaTrp-ValGlnPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300  
Db 543 AAGGCATATCCCTTACCTTCTGCTTCGCTCCGAGAGATGAAAGCCCTCGTTGAATG 602  
Qy 300 sLysAspArgTrnSerLeuTrpTyrMet-AlaValAsnLysAspLysSerTrpLysSer 320  
Db 603 CAAGCTTCTCCCTTACTTAATCTTACATGAGCTGTAAACAGAGAAAGATGATCTTCA 662  
Qy 320 snValGlyGlnTrpAspValAsnAlaValTrpTrp-GlyValPheProAlaLysGluIle 339  
Db 663 ACGTCAACCAACGACATCAACGCTGTGACATGGGGAGTTTCCCGACGTAAAGGAATT 722  
Qy 340 IleGln 341  
Db 723 ATACAA 728  
RESULT 14  
BE643238 821 bp mRNA linear EST 01-SEP-2000  
LOCUS BE643238  
DEFINITION Cri2\_8\_E11.SP6 Ceratopteris Spore Library Ceratopteris richardii  
cDNA clone Cri2\_8\_E11 5', mRNA sequence.  
ACCESSION BE643238  
VERSION BE643238.1 GI:9960923  
KEYWORDS EST.  
SOURCE Ceratopteris richardii.  
ORGANISM Ceratopteris richardii  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Filicophyta; Filicopsida; Filicales; Pteridaceae; Ceratopteris.  
REFERENCE 1 (bases 1 to 821)  
AUTHORS Chatterjee,A., San Miguel,P., Stout,S.C., Banks,J. and Roux,S.J.  
TITLE Expressed sequence tags of cDNA clones from a C. richardii library  
JOURNAL Unpublished (2000)  
COMMENT Contact: Roux SJ  
Section of Molecular Cell and Developmental Biology  
University of Texas  
Biology Building, Room 16, Austin, TX 78712, USA  
Tel: 512 471 4238  
Fax: 512 232 3402  
Email: [stroux@uts.cc.utexas.edu](mailto:stroux@uts.cc.utexas.edu)  
Plate: Cri2\_8 row: E column: 11  
Seq primer: Sp6.  
FEATURES  
source  
1..821  
/organism="Ceratopteris richardii"  
/cultivar="Brogn"  
/db\_xref="taxon:49495"

```

/clone="Cri2_8_E11"
/clone_lib="Ceratopteris Spore Library"
/tissue_type="Gametophyte"
/cell_type="Spore"
/dev_stage="20 hours after germination initiation"
/notes="Vector: pCMVSPORT6; EST sequence from cDNA library.
cDNA library constructed from mRNA isolated from C.
richardii spores that had developed for 20 hours after
their germination had been initiated by white light."
BASE COUNT      256 a      162 c      187 g      214 t
ORIGIN
Alignment Scores:
Pred. No.:      6,66e-117      Length:      821
Score:          1045.00      Matches:      188
Percent Similarity: 83.52%      Conservative: 40
Best Local Similarity: 68.86%      Mismatches: 45
Query Match:    47.78%      Indels:      0
DB:             10      Gaps:      0

US-09-720-451-6 (1-408) x BE643238 (1-821)
Qy      49 GlyphecyslysthrlyslleproalaasplilemetalaaleugluProilleysasp 68
      |||||||
      1 GGCTTTGCAAGACTAAGATTAAGTACTAGATTACAAATGCGTGAAGCTATCAAGAT 60
Qy      69 AsngluGLAAlaValysAlatryglyllehlsleuglythrcLmetCyslyslslylle 88
      |||||||
      61 GATGACAGAGCTGTACACATATGATATGATACCTGGGTACGAATGTGCAAAAAGATT 120
Qy      89 LeuAlaHisGlylleysthrleuHisleuYrThrlleuAsmetGluysSeraIaleu 108
      |||||||
      121 TTGGCTCATGGGTAAAGATCGCACTTTACACCTTAACCTGGAATAACACACTGTG 180
Qy      109 AlalleuemetksleuglyleuilegluJusertlysValSeraSserleuProtrp 128
      |||||||
      181 GGTATCTCAGAACCTTGCTGTAGATTATTCAAAATAGCGGGCACCTGCAATGCG 240
Qy      129 ArgArgProAlaasnValPheargValysGluasPvalArgProillePheArgAlaasn 148
      |||||||
      241 AGGCCTCCAGCAATATCAACAGAGATTAGAGAGATGTCGGCCATTTTCGGGCCAT 300
Qy      149 ArgProlysSertyrilSeraArgThrilegYrTPASPGLrTYrPROHISGLYArgTrp 168
      |||||||
      301 CGRCCCTGGAGTATACATCTCTAGACAGAACCTGGAGAGATTTCCTCGGGAAGTGG 360
Qy      169 GlyAspserCysasnProserTYrGlyAlaIeuserAspTYrGlnPheMetArgProArg 188
      |||||||
      361 GGGACACTCGCAACCTGATATGAGGCTTATACTGATTACCAAGTTCTACGCAGAGAG 420
Qy      189 AlaArgAspIysleuValgluJutrpAlaValProleuIysSeraValgluAspile 208
      |||||||
      421 TCTCGCAACAAGAAATCCAAAGAGGTGTCCTTGTGAATGCATCAACAGCAAT 480
Qy      209 TyrgluArgPheargleuYrCysleuglylysleuArgSeraSnProtrpSergluen 228
      |||||||
      481 TATCAGTATTTGCCAATATTCCTTCAGGGAATAAAGAACCAATCTTGGTCAAGGCT 540
Qy      229 AspIysleuGlnProgluThrllyslleleasnGluGluysSeraSeraSeraSera 248
      |||||||
      541 GAAGCTTGGAGCCAGCAAGAACTGAAATCAATCAATTTAGTAGAGATCAACTTAA 600
Qy      249 GlyPheleuthrIleasnSeraGlnProAlaValasnGlyIulysSeraSeraSera 268
      |||||||
      601 GGTTCCTCACAATTAATATGCTGCTGTTAATGGGAAATAACAGACTCTCCACA 660
Qy      269 ValGlyTrpGlyGlyProGlyGlyTYrValTYrGlnlysalatryValGlnPhePheCys 288
      |||||||
      661 TATGCTTGGGAGAGCTCTGCTGATATGTTACACAGAGCAATATTAATTTTCTCT 720
Qy      289 SerIysGluIysleuAspAlaValAspIysCysIysAspArgThrSerleuthrTYr 308
      |||||||
      721 TCACGGGCAAAACTTGACAAACTGTAGAGAAAGCAAAATCTTCCATCTTGACTTAT 780

```

```

Qy      309 MetAlaValAsnIysAspIysSertyrIysSeraSnVal 321
      |||||||
Db      781 ATTGCCGCTCACTCAAGGCTGAGACTCAGCAATATT 819

RESULT 15
BF071278
LOCUS
DEFINITION
BF071278          588 bp      mRNA      linear      EST 06-DEC-2001
s142e03.y1 Gm-c1067 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1067-1734.5, similar to SW:MTNR_ARATH O80585 PROBABLE
METHYLENETETRAHYDROFOLATE REDUCTASE ;, mRNA sequence.

ACCESSION
BF071278
VERSION
BF071278.1 GI:10846202
KEYWORDS
EST.
SOURCE
soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
1 (bases 1 to 588)
Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khana
,A., Bolla,B., Maira,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Materston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntersville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1779 Std Error: 0.00
High quality sequence stop: 414.

FEATURES
            source
            1..588
            /organism="Glycine max"
            /db_xref="taxon:3847"
            /clone="GENOME SYSTEMS CLONE ID: Gm-c1067-1734"
            /clone_lib="Gm-c1067"
            /tissue_type="germinating shoot, 3 day old seedling, auxin
            treatment"
            /lab_host="DH10B"
            /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
            XhoI; The cDNA library was constructed from mRNA isolated
            from germinating shoots of 3 day old seedling for the
            cultivar Williams 82. The seedlings were germinated in a
            growth chamber using germination paper in a solution
            containing 100ppm auxin. Complementary DNA was synthesized
            from mRNA using a primer consisting of a poly(GT) sequence
            with a XhoI restriction site. EcoRI adapters were ligated
            to the blunt-ended cDNA fragments followed by XhoI
            digestion. The cDNA fragments were directionally cloned
            into the EcoRI-XhoI restriction site of the pBluescript
            vector. The ligated cDNA fragments were transformed into
            DH10B host cells (GibcoBRL). This library was constructed
            in the laboratory of Dr. Randy Shoemaker."
BASE COUNT      182 a      114 c      124 g      168 t
ORIGIN
Alignment Scores:
Pred. No.:      8.81e-117      Length:      588
Score:          1042.00      Matches:      194
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.49%      Mismatches: 0
Query Match:    47.65%      Indels:      0

```

DB: 12 Gaps: 0

US-09-720-451-6 (1-408) x BF071278 (1-588)

OY 3 ValIleValIThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsnAspCys 22  
 ::

DB 4 CTGATTGTCCACCAATTTTATGATAGAGATATATCCATAATTTGATGACGACTGT 63  
 ::

OY 23 ArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsnTyrLys 42  
 ::

DB 64 CCCCAATTGGATATACCTGTCCATTGTACCTGGAATTATGCCCATTAATAATTACAG 123  
 ::

OY 43 GlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMetAla 62  
 ::

DB 124 GGCTTATCCGATGACGTGATTTGCAAAACAAAGATCCAGCTGACATTATGCTGCT 183  
 ::

OY 63 LeuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeuGlyThr 82  
 ::

DB 184 TTAGAGCCATTCAGACATGAGAGCTGTAGGAGCTTATGGATTCACCTGGAGACT 243  
 ::

OY 83 GluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThrLeuAsn 102  
 ::

DB 244 GAAATGTGCAAAAGATTTTACCTCATGGAATTAGACATTGCATCTTATACACTAAT 303  
 ::

OY 103 MetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSerLysVal 122  
 ::

DB 304 ATGGAGAAATCTGCATTGGCAATTAATAAGACCTTGCCCTAATTGAAGAGTCCAAAGTT 363  
 ::

OY 123 SerArgSerLeuProTyrArgArgProAlaAsnValPheArgValLysGluAspValArg 142  
 ::

DB 364 TCTAGTCTTACCTTGGAGAGCCCTGCAAATGTTTCCGTGTTAAAGAGATGCCGT 423  
 ::

OY 143 ProIlePheTyrAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTyrAspGln 162  
 ::

DB 424 CCAATCTTTTGGCAAAATGCACCAAAAGCTACATATCAAGGACCATGAGTGGATCAA 483  
 ::

OY 163 TyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyr 182  
 ::

DB 484 TACCCACATGGGCGGTGGGTGATTCCTGTAAATCCATCATATGTCATATCTGATTAT 543  
 ::

OY 183 GlnPheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197  
 ::

DB 544 CAGTTTCATCGGCGCACGTGCACGGGACAGAGCTTGTGAAGAA 588  
 ::

Search completed: February 15, 2003, 00:51:19  
 Job time : 2170 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2003, 21:53:11 ; Search time 273 Seconds  
(without alignments)  
3365.624 Million cell updates/sec

Title: US-09-720-451-6  
2187  
Sequence: 1 ARVITQLFYDTDFILKEVN.....VSLVDNDYINGDLFAVFAADF 408

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09720451/runat\_05022003\_074809\_2374/app\_query.fasta.1.583  
-DB=N\_Geneseq\_101002 -QWRT=fastap -SUFFIX=trng -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cd1  
-LIST=45 -DOCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09720451 @cgn.1.1.0 @runat.05022003\_074809\_2374 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEROUDRY -NEG\_SCORES=0 -MAIT -LONGLOG -DEV\_TIMEOUT=120  
-MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
N\_Geneseq\_101002: \*  
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1834	83.9	1821	21 AAC43311	Arabidopsis thalia
2	1834	83.9	1942	21 AAC49168	Arabidopsis thalia
3	1760	80.5	1335	21 AAC36967	Arabidopsis thalia
4	1184	54.1	1493	21 AAC44829	Arabidopsis thalia
5	998	45.6	558	21 AA250061	Soybean 5,10-methy
6	979	44.8	2187	22 AAC91206	Human methylenele
7	979	44.8	2219	17 AA709689	Human methylenele
8	979	44.8	2219	22 AAD20463	Human methylenele
9	979	44.8	2230	17 AA709694	Human methylenele
10	979	44.8	2230	21 AA550633	Human methylenele
11	979	44.8	2230	22 AAD20462	Human methylenele
12	979	44.8	2230	24 ABR13501	DNA encoding human
13	979	44.8	2230	24 ABR13539	Human methylenele
14	979	44.8	2230	24 ABR13540	Human methylenele
15	979	44.8	2230	24 ABR13541	Human methylenele
16	979	44.8	2230	24 ABR13545	Human methylenele
17	979	44.8	2230	24 ABR13546	Human methylenele
18	979	44.8	2230	24 ABR13550	Human methylenele
19	978	44.7	2230	24 ABR13542	Human methylenele
20	976	44.6	2230	24 ABR13547	Human methylenele
21	973	44.5	2230	24 ABR13543	Human methylenele
22	971	44.4	2230	24 ABR13548	Human methylenele
23	969	44.3	2230	24 ABR13544	Human methylenele
24	950	43.4	832	24 ABR98922	Arabidopsis thalia
25	742	33.9	3283	21 AA550634	Human methylenele
26	706.5	32.3	3181	21 AA550635	Mouse methylenele
27	415	19.0	23748	22 AA105529	Human reproductive
28	415	19.0	23748	22 AB198112	Human testicular a
29	386.5	17.7	451	21 AA250062	Wheat 5,10-methy
30	255	11.7	159	24 AB175279	Corn tassell-derive
31	188	8.6	2440	23 AA593090	DNA encoding novel
32	187.5	8.3	349980	21 AA597261	Neisseria meningit
33	182	8.3	876	24 AA597261	Neisseria meningit
34	172.5	7.9	283	22 AAD20482	Human MTRFR gene e
35	172.5	7.9	283	22 ABR13510	Human methylenele
36	169	7.7	235	22 AAD20480	Human MTRFR gene e
37	169	7.7	235	24 ABR13508	Human methylenele
38	168	7.7	640681	24 ABA92787	Buchnera sp. genom
39	166	7.6	235	22 AAD20491	Mouse MTRFR gene e
40	166	7.6	235	24 ABR13519	Mouse methylenele
41	160.5	7.3	283	22 AAD20493	Mouse MTRFR gene e
42	160.5	7.3	283	24 ABR13521	Mouse methylenele
43	158	7.2	978	22 AAH67344	C glutamicum codin
44	158	7.2	1005	22 AAF72089	Corynebacterium gl
45	158	7.2	349980	22 AAH68531	C glutamicum codin

## ALIGNMENTS

RESULT 1  
AAC43311  
ID AAC43311 standard; DNA; 1821 Bp.  
AC AAC43311;  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38803.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX EPI033405-A2.  
XX

PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148568.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.

```
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Alignment Scores:

Pred. No.:	1,55e-179	Length:	1821
Score:	1834.00	Matches:	334
Percent Similarity:	91.15%	Conservative:	37
Best Local Similarity:	82.06%	Mismatches:	36
Query Match:	83.86%	Indels:	0
DB:	21	Gaps:	0

US-09-720-451-6 (1-408) x AAC43311 (1-1821)

```
QY 1 AIAAGVAlIleValThGlnLeuPheTyRAspThrAspIlePheLeuLysPheValAsn 20
   ||| .....
DB 595 GGTGATCGATGTTGATCTACAGCTTTCTATGATGATGATATATCTCAAGTTGTGAT 654

QY 21 AspCysArgGlnIleGlyLleThrCysProIleValProGlyLleMetProIleAsnAsn 40
   ||| .....
DB 655 GATTGCGGCAAAATGGGATGATGTCCTCCATGTTCTCGAATATGCTATTAATTAAC 714

QY 41 TyrLysGlyPheLleArgMetThrGlyPheCysLysThrLysLleProAlaAspLleMet 60
   ||| .....
DB 715 TACAGAGGCTTTTGGTATGACTGTTCTCAAGACTAAAGATACCGAGTTGAGGTGATG 774

QY 61 AIAAlaLeuGluProIleLysAspAsnGluIuaIaValIaValAlaTyGlyLleHisLeu 80
   ||| .....
DB 775 GCTGCTTGACCTATCAAGATAAAGAGAGAGAGCTGTGAAGCTATGATTCACCTT 834

QY 81 GlyThrGluMetCysLysLysLleLeuAlaHisGlyLleLysThrLeuHisLeuTyrThr 100
   ||| .....
DB 835 GGAACAGAGATGTGTAAAAAGATGTGGCTCATGGAGTCAAGTCTCTCATCTTACACA 894

QY 101 LeuAsnMetGlnLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSer 120
   ||| .....
DB 895 TTGAACATGAGAAATCTGCTTGCAATATGATGAAATCTTGATGATGATGATGATGATG 954
```

```
QY 121 LysValSerArgSerLeuProTyrPArgPArgProAlaAsnValPheArgValLysGluAsp 140
   ||| .....
DB 955 AAAATTTTCGCTTCTTTACCTCGGAGACCCCTGCAATGTTTTCGACTAAGAGACAT 1014

QY 141 ValArgProIlePheThrPalaAsnArgProLysSerTyrLleSerArgThrLleGlyTyrP 160
   ||| .....
DB 1015 GTGGCGCCCATTTTCTGGGCAAAAGCTCCAAAGAGCTACATTTCTAGAACCAGGCGTGG 1074

QY 161 AspGlnTyrProHisGlyArgTyrPArgLysPserCysAsnProSerTyrGlyAlaLeuSer 180
   ||| .....
DB 1075 GAAGATTTCCACCAAGCGCGGGGGGTGATTCACGAGCTTCATATAGGTGACTCTCG 1134

QY 181 AspTyrGluPheMetArgProArgProAlaArgAspLysLysLeuValGluGluTyrPalaVal 200
   ||| .....
DB 1135 GATCATCAGTTCTTCACGCTCCGAGACCTGACCAAGAACCTTCAACAAAGATGGTTGTC 1194

QY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
   ||| .....
DB 1195 CCACGTAAAGGTGTCAGATATTCAGAGAAATTCAGAGACCTGCTTGAAACCTT 1254

QY 221 ArgSerAsnProTyrPserGluLeuAspGlyLeuGlnProGluThrLysLleAsnGlu 240
   ||| .....
DB 1255 AAAAGAGTCCCTGCTGCTGAATTAAGTAGGACTCCAGCCAGACACAAAGATATTAACGAG 1314

QY 241 GlnLeuGluLysLleAsnThrLysGlyPheLeuThrLleAsnSerGlnProAlaValAla 260
   ||| .....
DB 1315 CAACGTATAAAGTCAACATCCAAAGGCTTCTTGACCATCAATAGCCAAACATCAGTCAAC 1374

QY 261 GlyLysLysSerAspSerProThrValGlyTyrPArgLysGlyTyrValTyrGln 280
   ||| .....
DB 1375 GCCGAGAGATCGATGCCACCACTGTTGGATGGGAGGTCTGTGGATATATACCA 1434

QY 281 LysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaValAspLysCys 300
   ||| .....
DB 1435 AAGGCTTATCTAAGATCTTCTCTCTCAAGAGCAAAATTAATGACAGTGGTGAAGAAATGC 1494

QY 301 LysAspArgThrSerLeuThrTyrMetAlaValAlaAsnLysAspLysSerTyrLysSerAsn 320
   ||| .....
DB 1495 AAAGCTTGGCCATGCATCTTACTTACATGGCTGTGAACAAAGAGAACAGTGGTATGAAC 1554

QY 321 ValGlyGlnThrAspValAlaAsnAlaValThrTyrGlyValPheProAlaLysGluIleIle 340
   ||| .....
DB 1555 ACTGCCCAACCTGATGGAATGCTGTACCTTGGGAGATTTCCGGCTTAAGAAATCATTT 1614

QY 341 GlnProThrLleValAspProValSerPheAsnValTyrLysAspGluAlaPheGluIle 360
   ||| .....
DB 1615 CAACCAACCATTTGCGATCCCGCTAGCTTCAAGCTCGGAAGATGAAGATTTGAGACT 1674

QY 361 TrpSerArgLysTyrPalaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380
   ||| .....
DB 1675 TGGTCAAGAGAGCTGGCTACTTGTACCCAGAGCCGACCTTCCAGAAACTTGCTCGAG 1734

QY 381 GluValGlyGlySerHisPheLeuValSerLeuValAspAsnAspTyrLleAsnGlyAsp 400
   ||| .....
DB 1735 GAGGTGAAGAACACTCTATTTGTGTAGCTTGTGTGAGAACATTTACATGATGCTGAC 1794

QY 401 LeuPheAlaValPheAlaAsp 407
   ||| .....
DB 1795 ATATTCCGGGTGTTGCTGAT 1815

RESULT 2
AAC49168
ID AAC49168 standard; DNA; 1942 BP.
XX
AC AAC49168:
XX
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 60181.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
```

XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
PF 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125588.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130649.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134570.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145818.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149375.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152263.  
PR 10-SEP-1999; 99US-0153070.

```
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1.69e-179
Score: 1834.00
Percent Similarity: 91.15%
Best Local Similarity: 82.06%
Query Match: 83.86%
DB: 21
Gaps: 0
Indels: 0
Matches: 334
Conservative: 37
Mismatch: 36
```

US-09-720-451-6 (1-408) x AAC49168 (1-1942)

```
QY 1 ALaarGyAlIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 559 GCTGATCGATGATTGATACACGCTTTTCATGATGATATATATTCCTCAAGATTGTGAAT 618
QY 21 AspcysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 619 GATGTGCGCAAAATGGGATTAGTTGCCATTTGCTCGAAATTAATGCCATTAAATTAAC 678
QY 41 TyrIysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 679 TACAGAGCGCTTTTGGCGATGATGATCGTTTCGCAAGACTAAGATACCAAGTTGAGGTATG 738
QY 61 ALaAlaIleuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 80
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 739 GCTGCTTGAGAGCTATCAACGATACGAAAGACGCTGAAAGCCATGATGATACACTT 798
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLysLeuThr 100
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 799 GGAACAGAGATGTGTAAAAAGATGTGGCTCATGGAGTCAAGTCTCTTCATCTCTACACA 858
```

```
QY 101 LeuAsnMetCyluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluIleuSer 120
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 859 TTGAACATGAGAAATCTGCTCTTGCAATATGATGAATCTGTGTATGATGATGATGCTC 918
QY 121 LysValSerArgSerLeuProTrpArgProAlaAsnValPheArgValLysGluAsp 140
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 919 AAAATTTCGTCTTCTTACCGCTGAGACCGCTCGCAAAAGTCTTCCGATCTAAGGAAAGT 978
QY 141 ValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGlyTrp 160
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 979 GTGCGCCCATTTTCTGGGCAACCGTCCAAAGAGCTACATTTCTTGAAACCAAGGCGTGG 1038
QY 161 AspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1039 GAAGATTTCCCAACAGGCGGCTGGGTGATTCACGCGAGCTTCATATGTGTCACACTCTG 1098
QY 181 AspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluIleuTrpAlaVal 200
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1099 GATCATCAGTCTTCACGCTCCGCGAGCAGCTGACACAGACTTCACACAAATGGCTGTTC 1158
QY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1159 CCAGTGAAGAGTGTGAAGATATTCAGAGAAATTCAGAGAGCTGCGCTTGAAAACCTT 1218
QY 221 ArgSerAsnProTrpSerGluLeuAspGlyLeuGlnProGluThrLysIleIleAsnGlu 240
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1219 AAAAGCAGTCCGTGCTGTGATAGATGAGATCTCCACCCAGACAAAGATTAATTAACGAG 1278
QY 241 GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1279 CAAGTGTAAAGATCAACATCCCAAGAGCTTCTGACCATCAATAAGCAACATCAGTCAAC 1338
QY 261 GlyGluLysSerAspSerProThrValGlyTrpGlyGlyProGlyGlyTyrValLysGln 280
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1339 GCCAGAGATCTGATCCCAACTGCTGATGGGAGGCTGCTTGATGATATATACCA 1398
QY 281 LysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1399 AAGGCTTATCTAGAGTCTTCTGCTCAAGAGAAATTAATGACAGTGGTGAAGAAATTC 1458
QY 301 LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrpLysSerAsn 320
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1459 AAAGCTTTGGCATCGATTAATCATGCTGCTGGAACAAAGAGAACAGTGGGTATCGAAC 1518
QY 321 ValGlyGlnThrAspValAsnAlaValIleThrTrpGlyValPheProAlaLysGluIle 340
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1519 ACTGCCAAGCTGATGTGAATGCTGTAACCTGGGAGTCTTCCGCTAAGGAATTCATT 1578
QY 341 GlnProThrIleValAspProValSerPheAsnValTrpLysAspGluAlaPheGluIle 360
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1579 CAACCAACCTTGTGATCCGCTGACCTTCACAGCTGGAAGAGATGAAACATTTGAGACT 1638
QY 361 TrpSerArgGlyTrpAlaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1639 TGTCAAGAAAGCTGGGCTACTGATCTGACCAAGAGCCGCTTCCAGAACTTGTCTGAG 1698
QY 381 GluValGlyCysSerHisThrLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1699 GAGGTGAAGACAGCTACTATTGTTGTAAGCTTGTGGAGAACGATTATCAATCAATGTGAC 1758
QY 401 LeuPheAlaValPheAlaAsp 407
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1759 ATATTGCGGCTGTTCTGAT 1779
```

RESULT 3  
AAC36967  
ID AAC36967 standard; DNA; 1335 BP.  
XX  
AC AAC36967;  
XX  
DT 17-OCT-2000 (first entry)  
XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15706.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
PD  
XX  
XX 06-SEP-2000.  
PF  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142035.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.

```
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 4,48e-172
Score: 1760.00
Percent Similarity: 90.77%
Best Local Similarity: 82.05%
Query Match: 80.48%
DB: 21
Gaps: 0
```

US-09-720-451-6 (1-408) x AAC36967 (1-1335)

```
QY 18 PheValAsnAspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetPro 37
DB 1 TTTGCTGAAATGATGTCGCGAAATGGATTAATTTGCCATTTGCTCGTAATTAAGCCT 60
QY 38 IleAsnAspTyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAla 57
DB 61 ATTATTAACCTACACGAGGCTTCTTGCGATGACTGGTTTCTGCAAGACTAAGATACAGTT 120
QY 58 AspIleMetAlaAlaLeuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGly 77
DB 121 GAGGTGATGGCTGCTTGAGGCTATCAAGCATACAGAAACCTGTGAAGCCTATGCT 180
QY 78 IleHisLeuGlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHis 97
```

```
DB 181 ATTCACCTTGAACAAGATGTTGTAATAAAGATGTTGGCTCATGAGTCAAGTCTCTCAT 240
QY 98 LeuTyrThrLeuAsnMetCyluysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIle 117
DB 241 CTTTACACATTAACATGAGGAAATCTGCTTTGCAAAATTAATGATCAATCTTGATGATT 300
QY 118 GluGluSerLysValSerArgSerLeuProTyrPArgArgProAlaAsnValPheArgVal 137
DB 301 GATGAGTCCAAAATTTCTGTTCTTTTACCCTTGAGACGCCCTGCAAAATGTTTCCGTACT 360
QY 138 LysGluAspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThr 157
DB 361 AAGGAAGATGTGGCCCATTTCTGGGCAACCGTCCAAAAGCTACATTCTTGAAACC 420
QY 158 IleGlyTyrAspGlnTyrProHisGlyArgTyrPArgLysAspSerCysAsnProSerTyrGly 177
DB 421 AAGGCTGGGAAGACTTCCCAAGCGGTGGGTGATTCACGACGCTTCATATGCT 480
QY 178 AlaLeuSerAspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197
DB 481 GCACCTCGGATCATCAGTTCACGTCGCGACGACGACGACAAAGCTTCAACAAGAA 540
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
DB 541 TGGGTGTCCCACTGAAAGTGTTCAGATATTCAGAGAAATTCAGAGAGCTGCTGCTT 600
QY 218 GlyLysLeuArgSerAsnProTyrSerGluLeuAspGlyLeuGlnProGlyIleThrLysIle 237
DB 601 GGAACCTTAAGAGCAGTCCGTGTAATTAAGATGACCTCCACGACGACGACAAAGTAA 660
QY 238 IleAsnGluGlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnPro 257
DB 661 ATAAAGACCAACTGGTAAAGTCAACTCCAAAGCCTTTCACCATCAATACCAACCA 720
QY 258 AlaValAsnGlyGluLysSerAspSerProThrValIleTyrPglyLysProGlyLysTyr 277
DB 721 TCAGTCAACGCCGAGAGATCTGATCCCAACTGTGATGGAGGAGGTCCTGTTGATAT 780
QY 278 ValTyrGlnLysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuVal 297
DB 781 GTRATCCAAAGAGCTTATCTAGATCTTCTGCTCAAGGAAATTAATGATGCACTGGTG 840
QY 298 AspLysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTyr 317
DB 841 GAGAAATGCAAAAGCTTTGGCATCGATCTTACATGAGGCTGTCAACAAAGAGACAGTGG 900
QY 318 LysSerAsnValGlyGlnThrAspValAsnAlaValIleThrTyrGlyValPheProAlaLys 337
DB 901 GTATCGAACACTGTCCAAGCTGATTAATGCTGTAACTTGGGGTGTTCGCCGCTAAG 960
QY 338 GluIleIleGlnProThrIleValAspProValSerPheAsnValTyrLysAspGluAla 357
DB 961 GAAATCATTTCAACCAACTGTCATGCCGCTACCTTCAACGTCGTGGAAGAGTAAACCA 1020
QY 358 PheGluIleTyrSerArgGlyTyrPalaSerLeuTyrProGluAspGluAlaSerArgLys 377
DB 1021 TTTGAGACTGTGTCAGAAAGCTGGGCTAAGTGTACCCAGAAAGCCGACCTTCCAGAAAC 1080
QY 378 LeuValGluGluValGlyLysSerHisPheLeuValSerLeuValAspAsnAspTyrIle 397
DB 1081 TTGCTCGAGGAGGATGAAGACAGTACTATTGTAACCTTGTTGAGAACGATTAATC 1140
QY 398 AsnGlyAspLeuPheAlaValPheAlaAsp 407
DB 1141 AATGGGACATATTCGCGCTCTTGTGAT 1170

RESULT 4
AAC44829
ID AAC44829 standard; DNA; 1493 BP.
XX AAC44829;
AC
XX 18-OCT-2000 (first entry)
```

XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 44287.	PR	21-JUN-1999;	99US-0139817.
DE		PR	22-JUN-1999;	99US-0139899.
XX	Hybridisation assay; genetic mapping; gene expression control;	PR	23-JUN-1999;	99US-0140353.
KW	protein identification; signal transduction pathway;	PR	23-JUN-1999;	99US-0140354.
KW	metabolic pathway; promoter; termination sequence; ss.	PR	24-JUN-1999;	99US-0140695.
XX		PR	28-JUN-1999;	99US-0140823.
OS	Arabidopsis thaliana.	PR	29-JUN-1999;	99US-0140991.
PM	EP1033405-A2.	PR	30-JUN-1999;	99US-0141287.
XX		PR	01-JUL-1999;	99US-0141842.
XX		PR	01-JUL-1999;	99US-0142154.
PD	06-SEP-2000.	PR	02-JUL-1999;	99US-0142055.
XX		PR	06-JUL-1999;	99US-0142390.
PF	25-FEB-2000; 2000EP-0301439.	PR	08-JUL-1999;	99US-0142803.
XX		PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999;	PR	21-JUL-1999;	99US-0145088.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	07-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147028.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148365.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149920.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.

```

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160780.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:

Pred. No.:	1.89e-112	Length:	1493
Score:	1184.00	Matches:	213
Percent Similarity:	89.26%	Conservative:	28
Best Local Similarity:	78.89%	Mismatches:	29
Query Match:	54.14%	Indels:	0
DB:	21	Gaps:	0

US-09-720-451-6 (1-408) x AAC44829 (1-1493)

```

QY 1 ALaarYValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
DB 642 GCAGATTGATGTGACACACCTTTCTATGATGATGATATATCTCAAGTTGTGAAT 701
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
DB 702 GATTGTGCGCAATCGGATTAATGTGTCCTGGAATTAATGCTATTTCACAC 761
QY 41 TyrIysGlyPheIleLeuArgMetThrGlyPheCysLysThrIysIleProIleAspIleMet 60
DB 762 TACAAGGGGTTCTTGCGATAGCGTGTTTCGTAAAGACCAAGATACCCGCTGAGACTACT 821
QY 61 AlaIleuGluProIleLysAspAsnGluIleValIleLysAlaTyrGlyIleHisLeu 80

```

```

DB 822 GCTGCCTTAGACCTATTAGGATATGACGAGGCTGTAAAGCCTATGGAATCACTTT
QY 81 GlyThrGluMetCysLysIleLeuAlaHisGlyIleLeuThrLeuHisLeuYThr 100
DB 882 GCACAGAAATGTGCAAAAGATTTGGCCATGATATCTCCCTTATCTTACACA 941
QY 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluLys 120
DB 942 TTGAACGTGCAAAATCASCATATTGGATATTATACACCTTGCTGATTATAGTCA 1001
QY 121 LysValSerArgSerLeuProIleParArgArgProIleAsnValPheArgValIleGluAsp 140
DB 1002 AAATTTCTGCTCTCTACCTTGAGACGCCCTGCAAAATGTTTCGTTACTAAGAAAGAT 1061
QY 141 ValArgProIlePheThrPalaAsnArgProIleLysSerTyrIleSerArgThrIleGlyTyr 160
DB 1062 GTTCGCCCAATTTTGGGCCAAACCGTCCAAAGAGCTATATCTAAGAAAGGCGTGG 1121
QY 161 AspGlnTyrProHisArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180
DB 1122 AATGACTTCGCCACATGACGCTTGCGGTATTCACACAGTCGACATACGATCACTTTCG 1181
QY 181 AspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluIleTrpAlaVal 200
DB 1182 GATTATCAGTTTCGCGCCCAAAAGACGTGACAAAGACCTTCACAAAGATGGGTGTC 1241
QY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
DB 1242 CCACGTGAAAGCATTAAGATGTTCAAGAAATTCAAAGACCTGCGATTGGAACCTTA 1301
QY 221 ArgSerAsnProThrSerGlnLeuAspGlyLeuGlnProGlyIleThrLysIleLeuAsn 240
DB 1302 AAAAGCAGCCCATGCTCTGATATGATGACATCCAGCCAGACAAAGATCTAAATGAG 1361
QY 241 GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260
DB 1362 CAACCTGGAAAAATCAACGCCAAGGTTTCGTGACCATTAATAGCAACCATCACTGAT 1421
QY 261 GlyGluLysSerAspSerProThrValGly 270
DB 1422 GCAGCCAAATCCGATTCCTCAGCTATTGCT 1451

RESULT 5
AAZ50061
ID AAZ50061 standard; cDNA; 558 bp.
XX
AC AAZ50061;
XX
04-MAY-2000 (first entry)
DE Soybean 5,10-methylenetetrahydrofolate reductase cDNA clone.
XX
KW 5,10-methylenetetrahydrofolate reductase; MTHFR; soybean; inhibitor;
XX methionine synthesis; tetrahydrofolate metabolism enzyme; herbicide; ss.
OS Glycine max.
XX
FH key Location/Qualifiers
FT mat_peptide 3..557
FT /*tag_a
FT /partial
FT /product= "5,10-methylenetetrahydrofolate reductase"
FT /note= "encodes amino acids Ile4-Arg188 of the enzyme"
FT /EC_number= "1.7.99.5"
XX
PN WO200004163-A1.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15916.
XX
PR 15-JUL-1998; 98US-0092869.

```

XX (DUPO ) DU PONT DE NEMOURS & CO E. I.  
XX  
XX Falco SC, Fomodu LO;  
XX  
XX MPI; 2000-182429/16.  
XX P-PSDB; AAY44740.  
XX  
XX Novel tetrahydrofolate metabolism enzyme used to alter the level of  
PT tetrahydrofolate metabolism in plants and seeds  
XX  
XX  
XX Claim 3; Page 32; 37pp; English.  
XX  
XX The present sequence is a cDNA clone sfl1.pk0017.d12 encoding  
CC 5,10-methylenetetrahydrofolate reductase (MTHFR). The clone was  
CC isolated from a sfl1 cDNA library which was prepared using soybean  
CC immature flowers. MTHFR plays a role in the  
CC synthesis of methionine. The present sequence is used in the  
CC construction of a chimeric gene to alter the level of tetrahydrofolate  
CC metabolism enzymes in plants. The enzyme may provide target to facilitate  
CC design and/or identification of inhibitors that may be useful as  
CC herbicides. The polynucleotide is also useful as a source of probes for  
CC genetically and physically mapping the genes and as markers for traits  
CC linked to the genes.  
XX  
XX

SQ Sequence 558 BP; 170 A; 109 C; 116 G; 163 T; 0 other;

Alignment Scores:

Pred. No.: 8.07e-94 Length: 558  
Score: 998.00 Matches: 185  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 45.63% Indels: 0  
DB: 21 Gaps: 0

US-09-720-451-6 (1-408) x AAZ5061 (1-558)

QY 4 ILeValThrcGlnLeuPheThyrAspThrAspIlePheLeuLysPheValAsnAspCysArg 23  
DB 3 ATGTGTCACCAATTTATTTATGATACGAGATATTCCTCAAAATTTGACAGACTCGC 62  
QY 24 GlnIleGlyIleThrcysProIleValProGlyIleMetProIleAsnAsnIlyGly 43  
DB 63 CAATATGGAATACGTCGTCTATTTGACCGAATTAATGCGCCATTAATATTCACAGGCG 122  
QY 44 PheIleArgMetThrcGlyPheCysLysThrcLysIleProAlaLysIleMetAlaAlaLeu 63  
DB 123 TTTATCCGACGACTGGGTTTTCACAAACAAAGATACCAAGCTGATGAGCTGCTTA 182  
QY 64 GlnProIleLysAspAsnGlnGluAlaValAlaValAlaIleGlyIleIleIleGlyThrGlu 83  
DB 183 GAGCCTTTCAGACGACATGAGAAAGCTGTCAAGCCTTAATGAAATTAACCTGGAACTGAA 242  
QY 84 MetCysLysLysIleLeuAlaHisGlyIleLysThrcLysIleLeuLysIleLeuAsnMet 103  
DB 243 ATGGGAAAAAGATTTTACGTCATGGAATTAACACATTCGATTTTATACACATAATATG 302  
QY 104 GlnLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlnLysIleValSer 123  
DB 303 GAGAAATCTCATTTGGCAATCAATGAACCTTGCCCTAATGAAAGTCCAAAGTTTCT 362  
QY 124 ArgSerLeuProThrArgArgProAlaAsnValPheArgValAlaLysGlnAspValArgPro 143  
DB 363 AGGTCCTTACCTTGAGACGCCCTCGCAAAAGTTTCCGGTTAAAGATGTCCTCCA 422  
QY 144 IlePheThrAlaAsnArgProLysSerThrcIleSerArgThrIleGlyThrAspGlnIlyr 163  
DB 423 ATCTTTGGGCAATGACCAAAAAGCTACATATCAAGGACCATAGAGTGGGATCAATAC 482  
QY 164 ProHisGlyArgThrPheLysSerCysAsnProSerThrcIleValAlaLeuSerAspIlyrGln 183  
DB 483 CCACATGGGGTGGGGTGTATCTCGTAATCCATCATATGATGTCATTAATCTGATATATCAG 542

QY 184 PheMetArgProArg 188  
DB 543 TTGATGCGGCCACGT 557

RESULT 6

AAC91206  
ID AAC91206 standard; DNA; 2187 BP.

XX AAC91206;

XX 20-MAR-2001 (first entry)

XX Human methylenetetrahydrofolate reductase gene seq ID NO: 1.

XX Human: schizophrenia; developmental disorder; spina bifida cystica;  
XX Tourette's syndrome; bipolar illness; autism; conduct disorder;  
XX attention deficit disorder; obsessive compulsive disorder;  
XX chronic multiple tic syndrome; learning disorder; polymorphism; ds.  
XX  
XX OS Homo sapiens.  
XX  
XX WO200071754-A1.  
XX  
XX 30-NOV-2000.  
XX  
XX 24-MAY-2000; 2000WO-US14354.  
XX  
XX 25-MAY-1999; 99US-0318448.  
XX  
XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
XX  
XX Johnson WG, Stenroos ES;  
XX  
XX WPI; 2001-025174/03.  
XX  
XX

PT Diagnosing a developmental disorder, e.g. schizophrenia, by forming  
PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)  
PT and environmental variables affecting an individual and then comparing  
PT these DS with reference DS -  
XX  
XX

PS Disclosure: Page 114-115; 156pp; English.

XX The present invention provides a novel method of estimating the  
XX susceptibility of an individual to a developmental disorder using genetic  
XX and environmental variables. The method can be used in the diagnosis,  
XX prevention and treatment of disorders such as schizophrenia, spina bifida  
XX cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,  
XX attention deficit disorder, obsessive compulsive disorder, chronic  
XX multiple tic syndrome and learning disorders such as dyslexia.  
XX

SQ Sequence 2187 BP; 479 A; 655 C; 616 G; 437 T; 0 other;

Alignment Scores:

Pred. No.: 5.08e-91 Length: 2187  
Score: 979.00 Matches: 188  
Percent Similarity: 63.72% Conservative: 79  
Best Local Similarity: 44.87% Mismatches: 138  
Query Match: 44.76% Indels: 14  
DB: 22 Gaps: 7

US-09-720-451-6 (1-408) x AAC91206 (1-2187)

QY 1 AlaArgValIleValThrcGlnLeuPheThyrAspThrAspIlePheLeuLysPheValAsn 20  
DB 667 GCCGATTCATCATCACCACACCTTTCTTGTAGCCTGACACATCTTCCGCTTGTGAAG 726  
QY 21 AspCysArgGlnIleGlyIleThrcysProIleValProGlyIleMetProIleAsn 40  
DB 727 GCATGACCCGACATGGGATGATCATCTTGCCCATGCCCCGGGATCTTCCATCAGAGGC 786  
QY 41 TyrLysGlyPheIleArgMetThrcGlyPheCysLysThrcLysIleProAlaLysIleMet 60  
DB 787 TACCACTCCCTTGGGACACTTGTGAAGCTGTCCAAAGCTGGAGGTGCCACAGAAATCAAG 846

```

QY 61 AlaAlaLeuGluProIleuLeuAspAsnGluGluAlaValLeuAlaTyrGlyIleHisLeu 80
Db 847 GACGTGATTGAGCAATCAACAGACAGATGCTGCATCCGCAACTGATGCGACGAGCTG 906
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGly--IleLysThrLeuHisLeuTyr 99
Db 907 GCCGTGAGCCTGTGCGAGAGCTTCTGCGAGTGGCTGTGCGACAGGCCCTCCACTTCTAC 966
QY 100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlu 119
Db 967 ACCCTCAACCGCGAGATGGCTACACAGAGCTGCTGAAGCGCTGGGAGATGTGACTGTAG 1026
QY 120 SerLysValSerArgSerLeuProThrArgArgProAlaHisValPheArgValLysGlu 139
Db 1027 GACCCC--AGGCTCCCTTACCTGGGCTGTAGTGTCCACCCCAAGCCCGCAGAGAGAA 1083
QY 140 AspValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159
Db 1084 GATGTAGTCCCATCTTCTGGGCTCCAGACCAAGAGTTACATCTACGTTACCCAGGAG 1143
QY 160 TrpAspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
Db 1144 TGGGACGAGTCCCTTACAGCGCGCTGGGCAATTCCTTCCCTGCTTGGGAGAGCTG 1203
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197
Db 1204 AAGCACTACTACTCTTCTTCTACAGAGCAAGTCCCCCAAGGAGGAGTGTGTAGATG 1263
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1264 TGGGGGAGAGGAGCTGACGAGAGCAAGTGTCTTGAATCTTGTCTTACCTCTG 1323
QY 218 GlyLysLeuArgSerAsn-----ProTrpSerGluLeuAspGly 230
Db 1324 GGAGAACCAACCGCAATGGTCACAAAGTACTGTGCTGCCCTGGAGACGAT--GAGCCC 1380
QY 231 LeuGlnProGluThrLysIleIleAsnGluGluLeuGluLysIleAsnThrLysGlyPhe 250
Db 1381 CTGGCGGCTGAGACCGCTGCTGAGAGAGAGCTGCTGGGGGAGACCCCGAGGGCATC 1440
QY 251 LeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrValGly 270
Db 1441 CTCACCAATCAATCAACAGCCCAATCAAGGGAGCGTCCCTCCGACCCCATGCTGGC 1500
QY 271 TrpGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
Db 1501 TGGGGCCCGACGCGGGGCTATGCTCTCCAGAAAGCCCTACTAGATTTCACCTCCCGC 1560
QY 291 GluLysLeuAspAlaLeuValAspLysCysLysAsp--ArgThrSerLeuThrTyrMet 309
Db 1561 GAGACAGCGGAGACACTTCTGCAAGTGTGAGAGATACGAGTCCCGGGTTAATTACAC 1620
QY 310 AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
Db 1621 CTGTCAATGTGAAGAGGAAACATGCAACATGCCCTGAACTGACGCGAATCTGTCT 1680
QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
Db 1681 ACTTGGGCACTTCCCTGGGCGAGAGATCATTCACGCCACCGTAGTGATCCCGCTAGC 1740
QY 350 PheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgGlyTyrPalasLeuTyr 369
Db 1741 TTTCATGTTCTGGAAGAGAGAGGCTTGGCTGTGGATTTGAGCGGTGGGAAAGCTTAT 1800
QY 370 ProGluAspGluAlaSerArgLysLeuValGluGluValGlyLysSerHisPheLeuVal 389
Db 1801 GAGGAGGAGTCCCGCTCCGACCATCATCCAGTACATCCAGCAGACACTACTCTCTGTCT 1860
QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
Db 1861 AACCTGGTGGCAATGACTTCCACTGTGACAACTGCTTGGCAGGTGGTGGAGAC 1917

```

```

RESULT 7
AAT09689
ID AAT09689 standard; cDNA: 2219 BP.
XX
XX AAT09689;
AC
XX
XX 15-OCT-1996 (first entry)
DE Human methylene-tetrahydrofolate-reductase cDNA.
XX
XX Methylene-tetrahydrofolate-reductase; MTHFR; gene therapy;
XX cardiovacular disease; neurological disease; folic acid metabolism;
XX EC-1.5.1.20; enzyme; ss.
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 13..1980
XX FT /*tag= a
XX
XX WO9533054-A1.
XX
XX 07-DEC-1995.
XX
XX 25-MAY-1995; 95WO-CA00314.
XX
XX 26-MAY-1994; 94GB-0010620.
XX
XX (UYMC-) UNIV MCGILL.
XX
XX Goyette P, Rozen R;
XX
XX WPI: 1996-030565/03.
XX P-PSDB; AAR8358.
XX
XX Human methylene-tetrahydrofolate reductase cDNA probe - for
XX PT detection of sequence abnormalities in methylene-tetrahydrofolate
XX PT reductase e-g. in cardiovascular, neurological or folic acid
XX PT metabolism disorders
XX
XX PS Disclosure; Fig. 6A-6C; 66pp; English.
XX
XX
XX This sequence encoding human MTHFR has been localized to chromosome
XX CC 1p36.3. It may be used for the construction of DNA probes which may
XX CC be used for the identification of sequence abnormalities in patients
XX CC with severe or mild MTHFR deficiency. The resulting probe may also
XX CC be used in gene therapy to produce the MTHFR protein.
XX
XX SQ Sequence 2219 BP; 501 A; 656 C; 620 G; 442 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 5 19e-91 Length: 2219
XX Score: 979.00 Matches: 188
XX Percent Similarity: 63.72% Conservative: 79
XX Best Local Similarity: 44.87% Mismatches: 138
XX Query Match: 44.76% Indels: 14
XX DB: 17 Gaps: 7
XX
XX US-09-720-451-6 (1-408) x AAT09689 (1-2219)
QY 1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
Db 676 GCCGATTTCATCATACACGAGCTTTCTTGAAGCGTCAACATCTCTCCGCTTGTGAAG 735
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
Db 736 GCATGACGCGACATGAGGCACTTGTCCCATCTGCTCCCGGAGATCTTCCATCCAGGCG 795
QY 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaSpIleMet 60
Db 796 TACCACCTCCCTTGGCAGCTGTGAAGCTGTCCAGCTGTGAGGTGCGACAGAGATCAAG 855
QY 61 AlaAlaLeuGluProIleLysAspAsnGluGluAlaValLeuAlaTyrGlyIleHisLeu 80

```

```
Db      856 GAGGTGATTGAGCCAATCAAGACAGATGCTGCCATCGCAACTATGATGCAGCTG 915
      81 G1YThrgLumetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTy 99
      916 GCCGTGAGCTGTGCGCAGGAGCTTGTGGCCAGTGGCTTGTGCCAGGCGCTTCATC 975
Qy      100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlu 119
      976 ACCCTCAACCGGAGATGGCTACCCACAGAGTGCTCAACCGCGCTGGGAGATGTGAGCTG 1035
Qy      120 SerLysValSerArgSerLeuProTParArgProAlaAsnValPheArgValLysGlu 139
      1036 GACCCC---AGCGTCCCTACCTGCGCTCAGTGCACCCACCAAGGCGGAGAGGAA 1092
Qy      140 AspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGly 159
      1093 GATGTACGTCCCATCTTCTGGGCTCCAGACCAAGATTACATCTACCGTACCCAGAG 1152
Qy      160 TrpAspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
      1153 TGGGAGAGATGCCCTAACGGCGCTGGGCATTCCTCTCCCTGCGCTTGGGAGCTG 1212
Qy      180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu 197
      1213 AAGGACTACTACTCTTCTTACTGAGAGCAAGTCCGCCCAAGAGAGAGCTGCTGAAGATG 1272
Qy      198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
      1273 TGGGGGAGAGAGCTGACCCAGTGAAGCAAGTGTCTTGAAGTCTTGTCTTACTCTCG 1332
Qy      218 GlyLysLeuArgSerAsn-----ProTyrSerGluLeuAspGly 230
      1333 GGAGAACCAACCGGAATGGTCACAAGTACTGGCTGCGCTCGGAGACAT---GAGCCC 1389
Qy      231 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe 250
      1390 CTGGCGGCTGAGACCGCTCTGGAAGGAGAGCTGCGGGGTGAACCGGCAGCGGCATAC 1449
Qy      251 LeuThrIleAsnSerGlnProAlaValAlaAsnGlyLysSerAspSerProThrValGly 270
      1450 CTCACCATCAACTCAGAGCCCAACATCAACGGGAGCCGCTCCGACCCCATCTGGGCG 1509
Qy      271 TrpGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
      1510 TGGGGGCCAGCGGGGGCTATGCTTCCAGAGGCGCTAAGATTTCCTTCACTCCCGC 1569
Qy      291 GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
      1570 GAGACAGCGGAGACCTTCTGCAAGTGTGAAAGATACGAGCTCCGGGTTAATTACAC 1629
Qy      310 AlaValAsnLysAspGlySerTyrLysSerAsnValGlyGlnThrAspValAlaVal 329
      1630 CTGTGCAATGTAAGGGGTGAATAACATCAACCAATGCCCTGAACCTGACCCGAAATCTG 1689
Qy      330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
      1690 ACTTGGGCGATCTCTCGGGCAGAGATCAATCCAGCCACCGATGTGATCCGCTAGC 1749
Qy      350 PheAsnValTrpLysAspGluAlaPheGluIleTyrPserArgGlyTrpAlaSerLeuTyr 369
      1750 TTCATGTCTGTGGAGAGAGGCGCTTGTGCGTGTGATGAGCGGTGAGAAAGCTGTAT 1809
Qy      370 ProGluAspGluAlaSerArgLysLeuValGluValGlyGlySerHisPheLeuVal 389
      1810 GAGGAGAGAGTCCCGCTCCCGCAGCATATCCAGTACATCCACGACATCACTTCTGTGCTC 1869
Qy      390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
      1870 AACCTGTGTGACATGACTTCCCACTGACAAATGCTGTGAGGTGTGGAAGAC 1926
RESULT 8
AAD20463
```

```
ID      AAD20463 standard; cDNA; 2219 BP.
XX
AC      AAD20463:
XX
DT      03-JAN-2002 (first entry)
XX
DE      Human methylenetetrahydrofolate reductase (MTHFR) protein cDNA #2.
XX
KW      Human; methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;
KW      antisense therapy; EC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;
KW      pancreas; lung; brain; cytosolic; colon carcinoma; breast; colorectal;
KW      neuroblastoma; leukaemia; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      13..1983
FT      /*tag= a
FT      /product= "Human methylenetetrahydrofolate reductase
FT      (MTHFR) protein"
FT      /EC_number= "1.5.1.20"
XX
PN      US2001025030-A1.
XX
PD      27-SEP-2001.
XX
PF      01-DEC-2000; 2000US-0728910.
XX
PR      01-MAR-1999; 99US-0258928.
XX
PA      (ROZEN/) ROZEN R.
PA      (SEKH/) SEKHON J.
PI      Rozen R, Sekhon J:
XX
DR      WPI: 2001-638509/73.
DR      P-PSDB: AAE12607.
XX
PT      New antisense nucleic acids, which are methylenetetrahydrofolate
PT      reductase inhibitors, useful for treating, stabilizing or preventing
PT      cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or
PT      neuroblastoma
XX
PS      Disclosure; Fig 6; 68pp; English.
XX
CC      The invention relates to a non allele-specific antisense nucleic acids,
CC      which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)
CC      nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR
CC      catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to
CC      5-methylenetetrahydrofolate, a co-substrate for methylation of
CC      homocysteine to methionine. The invention provides potential therapy for
CC      individuals with MTHFR deficiency. The non allele-specific antisense
CC      nucleic acids are useful for treating, stabilizing or preventing cancer,
CC      particularly breast carcinoma, colon carcinoma, colorectal carcinoma,
CC      lung cancer, brain cancer, pancreatic cancer, kidney cell cancer,
CC      neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene
CC      therapy. The present sequence is human methylenetetrahydrofolate
CC      reductase (MTHFR) protein cDNA #2. Human MTHFR gene is mapped to
CC      chromosome 1p36.3.
XX
SQ      Sequence 2219 BP; 501 A; 657 C; 620 G; 441 T; 0 other:
XX
Alignment Scores:
Pred. No.: 5 19e-91 Length: 2219
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Indels: 14
DB: 22 Gaps: 7
US-09-720-451-6 (1-408) x AAD20463 (1-2219)
Qy      1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
```

```

Db 676 GCCGATTTCATCATCAGACGCTTTCTTGAGGCTGACACATTCTTCGGCTTGTGAAG 735
QY 21 ASPCARGSLINLEGLIETHRCYSPROILEVALPROGLYILEMETPLEASASN 40
Db 736 GCATGCACCGCATGGGCACTACCTGCCCCATGCTCCCGGAGATCTTCCATCCACAGGC 795
QY 41 TYRILYSGLYPHEILEARMETHRGLYPHECYSLYSTHRILYSILEPROLIAASILLMET 60
Db 796 TACCACTCCCTTGCGCAGCTGTGAAGCTGTCCAGCTGGAGGTGGAGGTCACAGAACAG 855
QY 61 ALAALAENGLUPROILELYSASPASNGLUVALAVALYSALATYRGLYLEHISLEU 80
Db 856 GACCTGATGTGACCAATCAAGACAGATGCTGCCATCCGCACTATGATGACAGACTG 915
QY 81 GLYTHRCUWETCYSLYSTLSILEUALAHISGLY---ILELYSTHRLHISLEUTYR 99
Db 916 GCGGTGACCTGTGCGAGGAGCTTGTGCCAGTGGCTGTGCCAGGCTCCACTTCTAC 975
QY 100 THRLEUASNMEGLULYSERIALALEUALALEUETALEUGLYLEU 119
Db 976 ACCCTCAACCGCAGATGGCTACACAGAGGTGTAAGCCGCTGGGATGTGAGACTGAG 1035
QY 120 SERLYSALSERISERLEUPROTPARAGARPROALIASNVALPHEARGVALYSGLU 139
Db 1036 GACCCC---AGGGCTCCCTACCTGGCTGCTCAGTGGCCACCCAGCCGAGAGGAA 1092
QY 140 ASPVALARGPROILEPHERPALAASNARGPROLYSSERTYRILESERARGTHRILEGLY 159
Db 1093 GATGTACGTCCCATCTCTGCGCTCCAGCAACCAAGATTCATCTACCCAGAGAG 1152
QY 160 TRPASPCLINTYRPROHISGLIARGTRPGLYASPSERCYSASNPROSETRYRGLIALALEU 179
Db 1153 TGGAGCAGATTCCCTAACGGCCGCTGGGCAATTCCTCTCCCTCCCTTTGGGAGACTG 1212
QY 180 SERASPTYRGLN-----PHEKETARGPROARGALARGASPLYSLSLEUVALGLU 197
Db 1213 AAGGACTACTACTCTCTCTCTGTAAGACAGCAAGTCCCCAAGAGAGAGACTGTAAGATG 1272
QY 198 TRPALAVALPROLEULYSERVALGLUASPILEYRGLIARGPHEARGLEUTYRCYSLEU 217
Db 1273 TGGGGGAGAGAGCTGACCACTGAGCAAGTGTCTTGAAAGTCTTTGCTTTACTCTCG 1332
QY 218 GLYLYSLEUARGSERASN-----PROTRPSERGLULEUASPGLY 230
Db 1333 GGAAGAACCAACCGGATGTGTCAAAAGTGAAGTCTGCTGCTGCTGGAAGAT---GAGCCC 1389
QY 231 LEUGLNPROGLIUTHRILYSILEILEASNGIUGILEUGLILYSILEASNTHRILYSGLYPHE 250
Db 1390 CTGGCGGCTGAGACCACTGCTGAAGGAGAGAGCTCTCGGGGTACACCGCCAGGCTATC 1449
QY 251 LEUTHRILEASNSEIRGLINPROALIAVALAASNGIYGLULYSERASPSERPROTHRYALGLY 270
Db 1450 CTCACCACTCAACTACAGCCCAACATCAACGGGAAGCCCTCCGACCATCGATGTGGGC 1509
QY 271 TRPGLYGLYPROGLIYTYRVALTYRGLINLYSALATYRVALGLIPHEPCYSEIRLYS 290
Db 1510 TGGGGCCCCACGCGGGGCTATGTCTTCAGAAAGCCCTACTTGAATTTTCTACCTCCGC 1559
QY 291 GLULYSLEUASPALAEUVALASPLYSYSLYSASP--ARGTHRSERLEUTHRYIWTET 309
Db 1570 GAGACAGCGGAGCACTTCTGCAAGTGTGAAGATGATGACGCTCCGGGTAAATTCACAC 1629
QY 310 ALAVALASNLYSASPGLYSERTYRPLYSERASNVALIGLILTHRASVALAVALAVAL 329
Db 1630 CTTGTCAATGTGAAGGTGAAACATCAACAGCCCTGAAAGCTGCACAGCCGAAGCTGTC 1689
QY 330 THRTPGLYVALPHEPROALALYSGLUILEILEGINPROTHRILEVALASPROVALSER 349
Db 1690 ACTTGGGAGCATCTTCCCTGGGAGAGATCAATCCAGCCACCGTAGTGATGCCGTGACG 1749
QY 350 PHEASNVALTRPPLYSASPGLUALAPHEGLUIETPSEARARGGLYTRPALASERLEUTYR 369

```

```

Db 1750 TTCATGTTCTGGAGAGACGAGSCCTTGGCCCTGTGATGTAGCGGTGGGAAAGCTGTAT 1809
QY 370 PROGLIUSPGLUALASERARGLYSLEUVALGLULIUALGLYGLYSERHISPLEUVAL 389
Db 1810 GAGGAGGAGTCCCGCTCCGACCATCATCTCCAGTACATCCAGACAACTACTCTGTGTC 1869
QY 390 SERLEUVALASPAASPTYR---ILEASNGIYASPLEUPHEALAVALPHEALASP 407
Db 1870 AACCTGTGGCAATGACTTCCCATGACACTGACACTGCTCTGGCAGGTGTGGAAGAC 1926

RESULT 9
AAT09694
ID AAT09694 standard; cDNA; 2220 BP.
AC AAT09694;
XX
XX
XX 15-OCT-1996 (first entry)
DE Human methylene-tetrahydrofolate-reductase cDNA.
KW Methylene-tetrahydrofolate-reductase; MTHFR; gene therapy;
KW cardiovascular disease; neurological disease; folic acid metabolism;
KW EC-1.5.1.20; enzyme; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FT mat_peptide 1..1980
XX FT /*tag= a
XX
XX W0533054-A1.
XX
XX 07-DEC-1995.
XX
XX 25-MAY-1995; 95WC-CA00314.
XX
XX 26-MAY-1994; 94GB-0010620.
XX
XX (UYWC-) UNIV MCGILL.
XX
XX PA
XX PI Goyette P, Rozen R;
XX
XX WPI: 1996-030565/03.
XX
XX P-PSDB: AAT09694.
XX
XX Human methylene-tetrahydrofolate reductase cDNA probe - for
XX detection of sequence abnormalities in methylene-tetrahydrofolate
XX reduction e.g. in cardiovascular, neurological or folic acid
XX metabolism disorders
XX
XX Disclosure: Fig. 1A-1F; 66pp; English.
XX
XX This sequence encoding human MTHFR has been localised to chromosome
XX 1p36.3. It may be used for the construction of DNA probes which may
XX CC be used for the identification of sequence abnormalities in patients
XX CC with severe or mild MTHFR deficiency. The resulting probe may also
XX be used in gene therapy to produce the MTHFR protein.
XX
XX
XX Sequence 2220 BP; 501 A; 657 C; 620 G; 442 T; 0 other;

Alignment Scores:
Pred. No.: 5,19e-91 Length: 2220
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Indels: 14
DB: 17 Gaps: 7

US-09-720-451-6 (1-408) x AAT09694 (1-2220)
QY 1 AIAARGVALILEVALTHGLINLEUPHETRYRASPTRHASPILLEPHEULYSPHEVALASN 20
Db 676 GCCGATTTCATCATCAGACGCTTTCTTGAGGCTGACACATTCTTCCGCTTGTGAAG 735

```

Qy	21	ASPCSAATgagInIleGGyIlethrcysProIleValProGlyIleMetProIleasn	40
Dy	736	GCATGGACCGCAATGGGATCATCTGCCCCATCGTCCCGGAACTTTCCCATCCAGGCG	795
Qy	41	TYPLyGlyPheIleArgMetThrGlyPheCysThrIleProIleAspIleMet	60
Dy	796	TACCACTCCCTTCGGCAAGCTTGGAAGCTGTCCAGCTGGAGGTGGCCACAGAGATCAAG	855
Qy	61	AlAlAlaLeuGlnProIleIleYsAspAsnGluGluAlaValIleSalIleIleHisLeu	80
Dy	856	GACGTGATTGACCAATCAAAAGCAACGATGTGGATCCGCAACTGTGCATCGAGCTG	915
Qy	81	GLYThrGluMetCysIleIleLeuAlaHisGly--IleIleThrLeuHisLeuIle	99
Dy	916	GGCTGAGCGCTGTGCAGAGAGCTTGTGGCAATGGCTGTGGCCAGGCGTCCACTTAC	975
Qy	100	ThrLeuAsnMetGluYsSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlu	119
Dy	976	ACCCCTAACCGCGAGATATGGCTACACAGAGGTGGTGAAGCGCGGGAGATGTGATGAG	1035
Qy	120	SerIleValSerIleSerLeuProIleProIleArgProIleAlaAsnValPheArgValysGlu	139
Dy	1036	GATGCC--AGCGTCCCTTCACCTGGGCTGTGATGGCCACCCCAAGCGCCAGAGAA	1092
Qy	140	AspValArgProIlePheThrPheAlaAsnArgProIlysserIleIleSerArgThrIleGly	159
Dy	1093	GAGTACGCTCCATCTTCTGGGCTCCAGACCAAGATGTACATCTACCTACCCAGAG	1152
Qy	160	TrpAspGlnIleProHisGlyArgTrpGlyAspSerCysAsnProSerIleGlyAlaLeu	179
Dy	1153	TGGGAGAGATTCCTTAAGGGCGGTGGGCAATTCCTTCCCTGCGCTTGGGAGCTG	1212
Qy	180	SerAspTrpGln-----PheMetArgProAlaAlaArgAspIleYsLeuAlaGlu	197
Dy	1213	AAAGACTACTACTCTTCTACTGAAAGACCAAGTCCCAAGAGGAGCTGCTGAAGATG	1272
Qy	198	TrpAlaValProLeuYsSerValGluAspIleIleTrpGluArgPheArgLeuIleYsLeu	217
Dy	1273	TGGGGGAGAGCACTGCACAGTGAAGCAAGTCTTTGAAAGCTTGTCTTACCTCTCG	1332
Qy	218	GLYLeuLeuArgSerAsn-----ProIleSerGluLeuAspGly	230
Dy	1333	GGAGAACCAACCGAATGGTGCACAAAGTGACTGCTGCGCTGTGAAGCAT--GAGGCC	1389
Qy	231	LeuGlnProGluThrIleYsIleIleAsnGluGlnLeuGlnIleAsnThrIleYsGlyPhe	250
Dy	1390	CTGGCGCGTGAAGACAGCTGTGTAAGAGAGAGCTGTGCGGGTGAAGCCGCAAGCATC	1449
Qy	251	LeuThrIleAsnSerGlnProAlaIleAsnGluIleYsLeuSerAspSerProThrValGly	270
Dy	1450	CTCCACATCACTCAGACCCCAACATCAACGGGAAGCGGTCTCTCCAGCCCATGTGGGC	1509
Qy	271	TrpGlyGlyProGlyGlyIleYsValIleYsGlnIleSalIleIleValGluPheCysSerIle	290
Dy	1510	TGGGGGCCCAAGCGGGGTGTGTCTTCCAGAAAGGCTCTACTTGAAGTTTTCATCTCCGC	1569
Qy	291	GluIleYsLeuAspAlaLeuValAlaPlyIleYsIleAsp--ArgThrSerIleThrIleYsIle	309
Dy	1570	GAGACAGCGGAAGCACTTCTGCAGAGTGTGAAGAGTACGAGCTCCGCGTATTATACAC	1629
Qy	310	AlaValAsnIleYsAspGlySerTrpIleYsSerAsnValGlyGlnThrAspValAsnAlaVal	329
Dy	1630	CTTGTCAATGTGAAGGTGAAGAAATCAACCAATGCGCCCTGAACGCAAGCGAATGTCTGC	1689
Qy	330	ThrTrpGlyValPheProAlaIleGluIleIleGlnProIleProIleValAspProValSer	349
Dy	1690	ACTTGGGAGATCTTCCCGGGCGAGAGATCATCCAGCCACCGTAATGTATCCCTGTACG	1749
Qy	350	PheAsnValIleTrpIleYsAspGluAlaIlePheGluIleIleTrpSerArgIleTrpAlaSerLeuTrp	369
Dy	1750	TTTATGTTCTTGTGAAGACGAGGCGTTTGGCTGTGTGATGTAAGCGGTGGGAAGCTGAT	1809

Accession	Gene	Chromosome	Position	Orientation	Length (bp)	Source	Notes
U00001	cdna	1	100000000	+	1000	GenBank	Human
U00002	cdna	1	100000000	+	1000	GenBank	Human
U00003	cdna	1	100000000	+	1000	GenBank	Human
U00004	cdna	1	100000000	+	1000	GenBank	Human
U00005	cdna	1	100000000	+	1000	GenBank	Human
U00006	cdna	1	100000000	+	1000	GenBank	Human
U00007	cdna	1	100000000	+	1000	GenBank	Human
U00008	cdna	1	100000000	+	1000	GenBank	Human
U00009	cdna	1	100000000	+	1000	GenBank	Human
U00010	cdna	1	100000000	+	1000	GenBank	Human
U00011	cdna	1	100000000	+	1000	GenBank	Human
U00012	cdna	1	100000000	+	1000	GenBank	Human
U00013	cdna	1	100000000	+	1000	GenBank	Human
U00014	cdna	1	100000000	+	1000	GenBank	Human
U00015	cdna	1	100000000	+	1000	GenBank	Human
U00016	cdna	1	100000000	+	1000	GenBank	Human
U00017	cdna	1	100000000	+	1000	GenBank	Human
U00018	cdna	1	100000000	+	1000	GenBank	Human
U00019	cdna	1	100000000	+	1000	GenBank	Human
U00020	cdna	1	100000000	+	1000	GenBank	Human
U00021	cdna	1	100000000	+	1000	GenBank	Human
U00022	cdna	1	100000000	+	1000	GenBank	Human
U00023	cdna	1	100000000	+	1000	GenBank	Human
U00024	cdna	1	100000000	+	1000	GenBank	Human
U00025	cdna	1	100000000	+	1000	GenBank	Human
U00026	cdna	1	100000000	+	1000	GenBank	Human
U00027	cdna	1	100000000	+	1000	GenBank	Human
U00028	cdna	1	100000000	+	1000	GenBank	Human
U00029	cdna	1	100000000	+	1000	GenBank	Human
U00030	cdna	1	100000000	+	1000	GenBank	Human
U00031	cdna	1	100000000	+	1000	GenBank	Human
U00032	cdna	1	100000000	+	1000	GenBank	Human
U00033	cdna	1	100000000	+	1000	GenBank	Human
U00034	cdna	1	100000000	+	1000	GenBank	Human
U00035	cdna	1	100000000	+	1000	GenBank	Human
U00036	cdna	1	100000000	+	1000	GenBank	Human
U00037	cdna	1	100000000	+	1000	GenBank	Human
U00038	cdna	1	100000000	+	1000	GenBank	Human
U00039	cdna	1	100000000	+	1000	GenBank	Human
U00040	cdna	1	100000000	+	1000	GenBank	Human
U00041	cdna	1	100000000	+	1000	GenBank	Human
U00042	cdna	1	100000000	+	1000	GenBank	Human
U00043	cdna	1	100000000	+	1000	GenBank	Human
U00044	cdna	1	100000000	+	1000	GenBank	Human
U00045	cdna	1	100000000	+	1000	GenBank	Human
U00046	cdna	1	100000000	+	1000	GenBank	Human
U00047	cdna	1	100000000	+	1000	GenBank	Human
U00048	cdna	1	100000000	+	1000	GenBank	Human
U00049	cdna	1	100000000	+	1000	GenBank	Human
U00050	cdna	1	100000000	+	1000	GenBank	Human
U00051							

XX  
XX Claim 1: Fig 1A-F; 93pp; English.

CC The present sequence is that of cDNA coding for human  
CC methylenetetrahydrofolate reductase (MTHFR, see AA196186), an enzyme  
CC catalysing the NADPH-linked reduction of 5,10-methylenetetrahydrofolate  
CC to 5-methyltetrahydrofolate, a co-substrate for methylation of  
CC homocysteine to methionine. To identify human MTHFR cDNA, PCR  
CC primers based on the porcine sequence were used to screen a human  
CC liver lambda-gt10 cDNA library by PCR. A 1266 bp fragment was  
CC obtained, and this was used to screen a human colon carcinoma cDNA  
CC library to obtain the 2.2 kb clone. The MTHFR gene (see AA05654)  
CC maps to chromosome 1p36.6. A cDNA probe for human MTHFR, which  
CC hybridises to the present sequence, is claimed. This probe can be  
CC used to identify MTHFR sequence abnormalities in individuals with  
CC severe or mild MTHFR deficiency. These abnormalities may comprise  
CC a mutation selected from 1676 to A, 4826 to A, 559C to T, 677C to  
CC T, 692C to T, 764C to T, 792+1G to A, 985C to T, 1015C to T, 1081C  
CC to T, 1298A to C and 1317T to C. MTHFR deficiency may be associated  
CC with a cardiovascular disorder, cancer (especially neuroblastoma or  
CC colorectal carcinoma), osteoporosis, neural tube defect in an  
CC offspring of a patient, neurological disorders, and other disorders  
CC influenced by folic acid metabolism. Also claimed are methods for  
CC treating MTHFR deficiency by gene therapy or by administration of  
CC MTHFR protein. Cancer can be treated by inhibiting MTHFR gene  
CC expression or MTHFR protein activity, or by administering an agent  
CC that modifies MTHFR gene expression.

XX  
SQ Sequence 2220 BP; 501 A; 658 C; 620 G; 441 T; 0 other;

Alignment Scores:  
Pridg. No.: 5,19e-91 Length: 2220  
Score: 979.00 Matches: 188  
Percent Similarity: 63.72% Conservative: 79  
Best Local Similarity: 44.87% Mismatches: 138  
Query Match: 44.76% Indels: 14  
DB: 21 Gaps: 7

US-09-720-451-6 (1-408) x AA050633 (1-2220)

QY 1 AAlaAGValIleValIleThGlnLeuPheTyrAspThrAspIlePheLeuValAsn 20  
DB 676 GCCGATTTCATCATCACCAGCTTTTCTTGAAGCTGACACATCTTCCGCTTGTGAAG 735  
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn 40  
DB 736 GCATGCACCGACATGGGATCATTGCCCATGCTCCCGGATCTTCCCATCCAGGCG 795  
QY 41 TyrIysGlyPheIleArgMetThrGlyPheCysIysThrIysIleProAlaSerIleMet 60  
DB 796 TACCACTCCCTTGGCAGACTTGTGAAGCTGTCCAAAGCTGGAGGCGCACAGAGATCAAG 855  
QY 61 AAlaAlaLeuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 80  
DB 856 GACGATGTTGAGCCAAATGAAGACACATGCTGCCATCCGCAATGTGGCATCGAGCTG 915  
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGly--IleLysThrLeuHisLeuTyr 99  
DB 916 GCGGTAGAGCTGTGCCAGAGACTTGTGGCAGTGTGGTGGCCAGAGCCCTCCACTGTAC 975  
QY 100 ThrIleAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnGluGlyIleGluGlu 119  
DB 976 ACCGTAAACCGCGAAGATGCTACACAGAGTGTGTGAAGCCCGTGGGAGATGTGAGTGA 1035  
QY 120 SerLysValSerArgSerLeuProThrArgArgProAlaAsnValPheArgValLysGlu 139  
DB 1036 GACCCC--AGGCTGCCCTTACGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGA 1092  
QY 140 AspValAlaArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159  
DB 1093 GATGTAAGTCCCATCTTGTGGGCTCCAGACCAAGATTCATCTACCCGATCCAGAG 1152  
QY 160 TrpAspGlnTyrProHisGlyAlaThrPheLysPheSerCysAsnProSerTyrGlyAlaLeu 179

DB 1153 TGGACAGAGTTCCCTTACAGCGCGTGGGCAATTCCTTCCCTCCCTTGGGAGTGG 1212  
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysIysLeuValGluGlu 197  
DB 1213 AAGCACTACTACTCTTCTTACCTGAAGAGCAATCCGCCAAGAGAGAGTGGTGAAGATG 1272  
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGlnArgPheArgLeuTyrCysLeu 217  
DB 1273 TGGGGGGGAGAGCTGACAGACAGAGAGCAAGATGCTTGAAGTCTTGTCTTACCTCTCG 1332  
QY 218 GlyLysLeuArgSerAsn-----ProTrpSerIleLeuAspGly 230  
DB 1333 GGAAGACCAACCGCAAGTGTCCAAAGTCACTGCTCCGCGAGACAT--GAGCCC 1389  
QY 221 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrIysGlyPhe 250  
DB 1390 CTGGGGGCTGAGACACCTGCTGAAGAGAGAGCTGTGGGGTGAACCGCAGGCGCATC 1449  
QY 251 LeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrValGly 270  
DB 1450 CTCACACATCACTCAGACGCCCAACATCAAGGAGAGCGCTCCGACCCCATCTGAGGC 1509  
QY 271 TrpGlyIleProGlyIleTyrValTyrGlnLysAlaIleTyrValGluPhePheCysSerIys 290  
DB 1510 TGGGGCCCGCAGCGGGGGCTATGCTTCCAGAGGCGCTACTAGATTCTTCACTCCCGC 1569  
QY 291 GluLysIleAsnAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309  
DB 1570 GAGACAGCGGAGAGCATCTTGTGCAAGTGTGAGAGTACAGAGCTCCGGGTAAATTCACAC 1629  
QY 310 AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329  
DB 1630 CTGTGCATGTGGAAGAGGTGAAGAAACATCAACCAATCCCTGAATGCGAGCGAATGTGTG 1689  
QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349  
DB 1690 ACTTGGGCACTCTCCCTGGGCGAGAGATCATCCAGCCACCGTAAGTGTCCCGTCAAC 1749  
QY 350 PheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgGlyTrpAlaSerLeuTyr 369  
DB 1750 TTCAATGTTCTGTGAAGAGAGAGGCTTTGCCCTGTGTGATGACGGGTGGGAAGCTGTAT 1809  
QY 370 ProGluAspGluAlaSerArgLysLeuValGluValGlyLysSerHisPheLeuVal 389  
DB 1810 GAGGAGAGTCCCGCTCCCGCACCATTCACATCCAGACACACAACTACTTCTGTGGTC 1869  
QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407  
DB 1870 AACCTGTGTGAACATGACTTCCCACTGTGCAACTGCTGTGGAGGTGTGGAAAGAC 1926

RESULT 11  
AAD20462  
ID AAD20462 standard; cDNA; 2220 BP.  
XX  
AC AAD20462;  
XX  
DT 03-JAN-2002 (first entry)  
XX  
DE Human methylenetetrahydrofolate reductase (MTHFR) protein cDNA #1.  
XX  
KW Human; methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;  
KW antisense therapy; BC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;  
KW pancreas; lung; brain; cytosolic; colon carcinoma; breast; colorectal;  
KW neuroblastoma; leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key location/Qualifiers  
FT 1..1963  
FT CDS /tag= a  
FT /product= "Human methylenetetrahydrofolate reductase  
FT (MTHFR) protein"









QY	61	AlaAlaLeuGluProIleuLysAspAsnGluGluAlaValAlaLysAlaIleIleHisLeu	80
Db	836	GACGTAGTGGCCCAATCAAGAACAAAGAGTGCCTCATCCGGAACATATGGATCCAGCTG	915
QY	81	GlyThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuIyr	99
Db	916	GCCCGGACCCGTGTCCAGAGAGCTTCTGGCAGTGGGCTTGTCGACAGCCTCCACTTCTAC	975
QY	100	ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGluIleIleGlu	119
Db	976	ACCCTCAACCCCGAGGAGGTACACAGAGAGTCTCAACGCGCTGGGATGTGGACTGAG	1035
QY	120	SerLysValSerAspSerLeuProIlePargArgProAlaAsnValPheArgAlaLysGlu	139
Db	1036	GACCC---AGCGTCCCTTACCTGGCGCTCTAGTGGCCACCACCCCAAGCGCGGAAGAA	10922
QY	140	AspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGly	159
Db	1093	GATGACGTCCCATCTCTTGGCGCTCCAGACCAAGAAGTATCATCTACCGTACCCAGAG	11522
QY	160	TrpAspGlnTyrProHisGlyArgTyrPheLysAspSerCysAsnProSerTyrGlyAlaLeu	179
Db	1133	TGGGACGAGTCCCTTACAGCGCCCTGGGGAAATTCCTCTCCCGCTTGGGGAGCTG	1212
QY	180	SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu	197
Db	1213	AAGGACTACTACGCTCTTACTACGTGAAGAGCAATCCCCCAAGAGAGAGCTGTGAAGATG	1272
QY	138	TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu	217
Db	1273	TGGGGGAGGAGCTGACGACGACGAGCAAGATGCTTGAAGTCTTGTTACTTCACTCTCG	13322
QY	218	GlyLysLeuArgSerAsn-----ProIleSerGluLeuAspLys	230
Db	1333	GGAAGACCAACCGGATGTGTCAACAAAGTACTTGCTCTCCCTGGAGACAT---GAGCC	1389
QY	231	LeuGlnProGluThrLysIleIleAsnGluGluLeuGluLysIleAsnThrLysGlyPhe	250
Db	1390	CTGGCGGCTGAGACCACTGCTGTGAAGAGAGACCTGCTGGGTGACCGCAGGCGCATC	1449
QY	251	LeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspSerProThrValGly	270
Db	1450	CTCAACATCACTACAGCCCAATCAATCAAGGGAACCCCTCTCCGACCCCATCGTGGC	1509
QY	271	TrpIleGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPheCysSerLys	290
Db	1510	TGGGGCCCAAGCGGGGCTATGTCTTCCAGAGAGCCTACTTGAATTTTCACTTCCGC	1569
QY	291	GluLysLeuAspAlaLeuValAspLysCysLysAsp--ArgThrSerLeuThrTyrMet	309
Db	1570	GAGACACCGGAAGACACTTGCACAGTGTGGAAGAGTACAGACTCCGGGTAAATTACAC	1629
QY	310	AlaValAsnLysAspGlySerTyrLysSerAsnValGlyGlnThrAspValAlaVal	329
Db	1630	CTTGTCAATGTGAAGGTTGAATAAACATCAACCAATGCCCTGAACTGCACCCGAAATGCTGT	1689
QY	330	ThrTrpIleValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer	349
Db	1690	ACTTGGGGGACTCTTCCCTGGCGAGAGATCATCCAGCCACCGATGTGATCCCGTACG	1749
QY	350	PheAsnValIleTyrLysAspGluAlaPheGluIleTyrPheSerArgGlyTyrPalaSerLeuTyr	369
Db	1750	TTTCAATGTTCTGGAAAGGACGAGGCTTGTGCCCTGTGGATTTGACCGGTGGGAAACCTGAT	1809
QY	370	ProGluAsnLysAlaSerArgLysLeuValGluGluValGlyLysSerHisPheLeuVal	389
Db	1810	GAGAGAGAGTCCCGGTCGCAACATCATTCAGTACATGCAAGCAACTACTTCTGTGTC	1869
QY	390	SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp	407
Db	1870	AACCTGTGGACATGACTTCCACATGTGCAACATGCTCTGTGCAAGGTGTGGAAAGAC	1926

ABK13541	ABK13541 standard; cDNA; 2220 BP.
ID	ABK13541
XX	
AC	ABK13541;
XX	
DT	09-APR-2002 (first entry)
XX	
DE	Human methylenetetrahydrofolate reductase (MTHFR) C559T allele.
XX	
KW	Methylenetetrahydrofolate reductase; MTHFR; neuroleptic;
KW	neuroprotective; tranquiliser; nootropic; antidepressant;
KW	anticonvulsant; cerebroprotective; amyotrophic lateral sclerosis;
KW	anxiety; dementia; depression; epilepsy; Huntington's disease;
KW	migraine; demyelinating disease; multiple sclerosis; pain;
KW	Parkinson's disease; psychosis; stroke; cardiovascular disorder;
KW	inborn error of metabolism; metabolic disease; endocrine disease;
KW	lancet error of metabolism; inflammation; immune disorder; human;
KW	neuroblastoma; colorectal carcinoma; neoplastic disease; renal disease;
ss	EC number 1.5.1.20; mutant.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
XX	Key
XX	Location/Qualifiers
FT	1..1983
FT	/*tag- a
FT	/product= "MTHFR"
FT	/note= "Methylenetetrahydrofolate reductase"
FT	replace(559,C)
FT	/*tag- b
XX	
XX	MO200196598-A2.
PN	
PD	20-DEC-2001.
XX	
XX	12-JUN-2001; 2001MO-CA00867.
PR	12-JUN-2000; 2000US-0592595.
XX	
PA	(UYMC-) UNIV MCGILL.
XX	
PI	Rozen R;
DR	WPI: 2002-130741/17.
DR	P-PSDB; AA075423.
XX	
PT	Diagnosing subjects at risk for or suffering from a psychosis,
PT	particularly schizophrenia comprises determining the presence of a
PT	heterozygous methylenetetrahydrofolate reductase mutant allele in the
XX	subject
PS	
PS	Claim 4; Page -: 102pp; English.
XX	
XX	The invention describes a method of diagnosing a psychosis in a subject,
XX	or a risk for or propensity to psychosis in a subject comprising
CC	determining the presence of a heterozygous methylenetetrahydrofolate
CC	reductase (MTHFR) mutant allele in a nucleic acid sample obtained from a
CC	subject. The method is useful for diagnosing subjects at risk of, or
CC	suffering from a psychosis, particularly schizophrenia but includes
CC	manic-depressive disease, organic psychotic disorders, psychosis in
CC	alcohol or drug intoxication, postinfection psychosis, postpartum
CC	psychosis, senile psychosis, traumatic psychosis and acute idiopathic
CC	psychotic illnesses. A new pharmaceutical composition is used for
CC	treating amyotrophic lateral sclerosis, anxiety, dementia, depression,
CC	epilepsy, Huntington's disease, migraine, demyelinating disease, multiple
CC	sclerosis, pain, Parkinson's disease, schizophrenia, psychoses, or
CC	stroke. Deficiency may be associated with diseases including
CC	cardiovascular disorders, cancer (e.g. neuroblastoma and colorectal
CC	carcinoma), osteoporosis, metabolic or endocrine disease, inborn errors
CC	of metabolism, inflammation, immune disorders, neoplastic disease and
CC	renal disease. This sequence encodes a mutant human
CC	methylenetetrahydrofolate reductase (MTHFR, EC number 1.5.1.20), in which
CC	nucleotide 559 has been altered to produce a mutant allele, described in

CC the method of the invention.  
 CC Note: This sequence does not appear in the specification but has been  
 CC created from the wild type sequence (ABK13501) using information given  
 CC in claim 4 of the invention.

XX Sequence 2220 BP; 501 A; 657 C; 620 G; 442 T; 0 other;

# Alignment Scores:

Pred. No.:	5,19e-91	Length:	2220
Score:	979.00	Matches:	188
Percent Similarity:	63.72%	Conservative:	79
Best Local Similarity:	44.87%	Mismatches:	138
Query Match:	44.76%	Indels:	14
Db:	24	Gaps:	7

US-09-720-451-6 (1-408) x ABK13541 (1-2220)

```

QY 1 AlaATgValIleValThGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
Db 676 GCCGATTTCATCATCATCCGACCTTTCTTGTAGGCTGACACATCTCTCCGCTTGTGAG 735
QY 21 AspCysArgInIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
Db 736 GCATCCACCGACACATGGGATCATCTTCCCATCGTCCCGGATCTTCCCATCCAGGGC 795
QY 41 TyrIlySglyPheIleArgMetThrGlyPheCysIlyThrIlyIleProAlaAspIleMet 60
Db 796 TACCACTCCCTTCGCGACCTTGTGAGCTGTCCAAAGCTGAGGAGGAGGCCACAGAAATCAG 855
QY 61 AlaAlaIleuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 80
Db 856 GACGCGATTGACCATTAAGACACATGCTGCTCCATCCCACTATGAGATCGAGCTG 915
QY 81 GlyThrIleMetCysLysIlyIleLeuAlaHisGly--IleLysThrIleuHisLeuTyr 99
Db 916 GCCGTGACCGTGTGCGAGACCTTCTGGCCAGTGGCTTGGCCAGGCTTCCACTTCTAC 975
QY 100 ThrIleuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlu 119
Db 976 ACCCTCAACCGGAGATGGCTACCAAGAGGCTGCTGAAGCGCTGGGAGATGTGAGCTGAG 1035
QY 120 SerLysValSerArgSerLeuProIlePheArgProAlaAsnValPheArgValLysGlu 139
Db 1036 GACCC---AGCGTCCCTTACCTGAGGCTCTCAGTGCACCCCAAGCGCCGAGAGAA 1092
QY 140 AspValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159
Db 1093 GATGTACGTCCCATCTTCTGGGCTCCAGACCAAGAGTTACATCTACCTACCCAGAGAG 1152
QY 160 TrpAspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
Db 1153 TGGGACGAGTCCCTTAACGGCGCTGGGCAATCTCTTCCCTGCTTGGGAGACTG 1212
QY 180 SerAspTyrGln-----PheMetArgProAlaArgAspLysLysLeuValGlu 197
Db 1213 AAGGACTACTACCTCTTCTTACCTGAAGCAAGTCCCAAGGAGGAGCTGTGAAGATG 1272
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1273 TGGGGGAGGAGCTGACCAAGCAAGGCTTTTGAAGTCTTGTCTTACCTCTG 1332
QY 218 GlyLysLeuArgSerAsn-----ProTrpSerGluLeuAspGly 230
Db 1333 GGAACAACCAACCGAATGGTCAACAAGTACTGCTGCCCTGGAACGAT--GAGCC 1389
QY 231 LeuGlnProGluThrLysIleIleAsnGluGlnLeuLysIleAsnThrLysGlyPhe 250
Db 1390 CTGGCGGCTGAGACACACCTGCTGAAGAGAGAGCTGCTGGGTGAACCGCAGGCGATC 1449
QY 251 LeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrValGly 270
Db 1450 CTCACATCAACTACAGCCCAACATCAAGGGAAGCGCTCTCCGACCCCATCGTGGGC 1509

```

```

QY 271 TrpGlyLysProGlyIlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
Db 1510 TGGGGCCCCAGCGGGGCTATGCTTCCAGAAAGCCCTACTTACAGTTTCCACTCCCG 1569
QY 291 GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
Db 1570 GAGACAGCGAAGCACTTGTGCAAGTCTGAAAGATGACGAGCTCCGGTTAATTACAC 1629
QY 310 AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
Db 1630 CTGTCAATGTGAAGGTGAAGAAACATCCAAATGCCCTGACTGACCGGAATGCTGTTC 1689
QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
Db 1690 ACTTGGGCGATCTTCCCTGCGGCGAGATCATCCAGCCACCGTAGTGATCCCGTCAGC 1749
QY 350 PheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgGlyTrpAlaSerLeuTyr 369
Db 1750 TTCATGTCTGGAAGGACGAGGCTTGGCTTGGATTTGAGCGGTGGGAAAGCTGTAT 1809
QY 370 ProGluAspGluAlaSerArgLysLeuValGluGluValGlyLysSerHisPheLeuVal 389
Db 1810 GAGGAGAGATCCCCGTCGCCACCATCATCCAGTACATCCAGACAACTACTTCTGTC 1869
QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
Db 1870 AACCTGTGACATATGACTTCCCATCTGACAACTGCTGTGGAGGTGGTGAAGAC 1926

```

Search completed: February 14, 2003, 23:16:00  
 Job time : 301 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2003, 21:20:31 ; Search time 41 Seconds

(without alignments)  
956.655 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187

Sequence: 1 ARVIYQLFYDIDIFLKFVN.....VSLVDNDYINGDLFAVADF 408

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1834	83.9	606	2	probable methylene
2	1777	81.3	592	2	methylenetetrahydr
3	895	40.9	603	2	methylenetetrahydr
4	778	35.6	599	2	methylenetetrahydr
5	711	32.5	615	2	hypothetical prote
6	629	28.8	641	2	methylenetetrahydr
7	578.5	26.5	657	2	methylenetetrahydr
8	486	22.2	416	2	methylenetetrahydr
9	421.5	19.3	566	2	methylenetetrahydr
10	222	10.2	290	2	methylenetetrahydr
11	222	10.2	296	2	methylenetetrahydr
12	221	10.1	314	2	methylenetetrahydr
13	215	9.8	275	2	methylenetetrahydr
14	193	8.8	307	2	methylenetetrahydr
15	189	8.6	292	1	methylenetetrahydr
16	189	8.6	296	1	methylenetetrahydr
17	189	8.6	296	2	methylenetetrahydr
18	189	8.6	296	2	methylenetetrahydr
19	189	8.6	296	2	methylenetetrahydr
20	188	8.6	282	2	methylenetetrahydr
21	184	8.4	296	1	methylenetetrahydr
22	182	8.3	292	2	methylenetetrahydr
23	182	8.3	292	2	methylenetetrahydr
24	182	8.3	292	2	methylenetetrahydr
25	179	8.2	294	2	methylenetetrahydr
26	179	8.2	296	2	methylenetetrahydr
27	179	8.2	312	2	methylenetetrahydr
28	171	7.8	304	2	methylenetetrahydr
29	168	7.7	292	2	methylenetetrahydr
29	163	7.5	283	2	hypothetical prote

## ALIGNMENTS

### RESULT 1

T00696

probable methylenetetrahydrofolate reductase [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein F6E13.29

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text-change 23-Mar-2001

C/Accession: T00696; B84875

R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke submitted to the EMBL Data Library, June 1998

A:Description: Arabidopsis thaliana chromosome II BAC F6E13 genomic sequence.

A:Reference number: Z14180

A:Accession: T00696

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-606 <RNU>

A/Cross-references: EMBL:AC004005; NID:g3212846; PID:g3212869

A/Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; M01D:20083487; PMID:10617197

A:Accession: B84875

A>Status: preliminary

A:Molecule type: DNA

A/Residues: 1-606 <STO>

A/Cross-references: GB:AE002093; NID:g3212869; PIDN:MAC23420.1; GSPDB:GN00139

C/Genetics:

A:Gene: At2g44160; F6E13.29

A:Map position: 2

A:Insertions: 155/1; 194/3; 252/3; 310/3; 344/2; 381/3; 408/3; 468/1; 579/3

Query Match

Best Local Similarity 83.9%; Score 1834; DB 2; Length 606;

Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

QY	1	ARVIYQLFYDIDIFLKFVNDCRQIGITPCIVGIMPINNNGKIFMTGFCRTKIPADIM	60
DB	199	ADIVITQLFYDIDIFLKFVNDCRQIGISCPVIGIMPINNNGKIFMTGFCRTKIPVEVA	258
QY	61	AALPEIKDNEAVKAGIHGTECMCKIIAHGKTHLTKTNMNSALAIIMNLGITEES	120
DB	259	AALPEIKDNEAVKAGIHGTECMCKIIAHGKTHLTKTNMNSALAIIMNLGITEES	318
QY	121	KVSRSLEPWRPVPANFVKEDVPRIFWANRPKSYISRTIGMDQYPHGRWDCSNPSYGALS	180
DB	319	KISRSLEPWRPVPANFVKEDVPRIFWANRPKSYISRTIGMDQYPHGRWDCSNPSYGALS	378
QY	181	DYGFMPRRARDDKLVEMAVPLKSVSDIYERRRLYCLGLRNSPSELDGLQPETRIINE	240
DB	379	DHGFSPRRARDDKLVQEWVPLKSVSDIYERRRLYCLGLRNSPSELDGLQPETRIINE	438

OY 241 QLEKINTKGFLTINSQPAVNGEKSDSPVTGNGGGCGGYQQAKAYEFPCKSKLDALVDKC 300  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 439 QLIVNNSKGFLTINSQPVSNAERNRSDSPVTGWGPVGYYQAYLEFFCSKERLDAVYEKC 498  
OY 301 KDRSLTFMAVNVNKDGSWKSNNVGOTDVNAVTVGTVPFAKEIIPTIYDPVPSFVNMDAEFEI 360  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 499 KALSITITMAYNVKKGGQMVNTAQADVNAVTVGTGVFAKEIITPTIYDPASFSVMNDDEARET 558  
OY 361 WSRGNASLIYPEDNASRKLYEEVGGSHFLVSLVDNDYINGDLFAAFAD 407  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 559 WSRSMANLVPEADPSRNLLLEEYKNXYLSLVENDYINGDLFAAFAD 605

RESULT 2  
T47821  
methylinetetrahydrofolate reductase MTHFR1 - Arabidopsis thaliana  
N:Alternate names: protein F24G16.240  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47821  
R:D:Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.,  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24477  
A:Accession: T47821  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-592 <DNA>  
A:Cross-references: EMBL:AL138647  
A:Experimental source: cultivar Columbia; BAC clone F24G16  
C:Genetics:  
A:Map position: 3  
A:Introns: 155/1; 182/3; 240/3; 298/3; 332/2; 369/3; 396/3; 456/1; 567/3  
A>Note: F24G16.240

[illegible]

```

RESULT 3
138920
methylentetrahydrofolate reductase 2 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

```

C:Accession: T38920  
R:Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajadream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: Z21817  
A:Accession: T38920  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-603 <P>A>  
A:Cross-references: EMBL:Z69728; PIDN:CA93581.1; GSPDB:GN00066; SPDB:SPAC56F8.10  
A:Experimental source: strain 972h-; cosmid c56F8  
C:Genetics:  
A:Gene: SPDB:SPAC56F8.10  
A:Map position: 1

```

Query Match Similarity      40.9%: Score 895; DB 2: Length 603;
Best Local Similarity      45.2%: Pred. No. 1.6e-63;
Matches 189; Conservative  57; Mismatches 148; Indels  24; Gaps  8;

QY      1 ARVITQLEFYDTDFLEKRVNDRCROIGTCPTVPGPIMPINNYKGFIRMTGCFKTKIPADIM 60
      1 | | | | : | | | : | : | : | | | | | | : | | | | | | | | |
DB      178 ADFITQGFIEYVDNRIANVADKYRAAGINIPFPGIMP IQADSPDIRAKMSGVIRPQHFM 237
      1 | | | | : | | | : | : | : | | | | | | : | | | | | | | | |

QY      61 AALEPIKONEEAAYKAYGIHLCGTEMCCKTLANGITLHYLTINMKRSALATIMNGLLEES 120
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |
DB      238 DTLVAVKKDDDEEYRBERGVELIVEMCKRLIAGIGIRLHFHYTMNLEAKYMIIERGLDEN 297
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |

QY      121 KYSRSLPMRRPANY-----PFVKEDVGRPIFMANPKYSIRTIQMDQYPHGRMGDS 171
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |
DB      298 LA----PIYDINNVELTIANSSODRININGVPIFWRIRNYESYRTOQMDELPHGRMGDS 353
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |

QY      172 CNPSYGALSDYQF--MRPRADKRLVEEMVPLKSVEDIYERFLYCLGKLNSPMSSEL 228
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |
DB      354 RSPARGEDFAYRYGRMSP---KEITTSWSP--KSYSEIIDLPAFYCEKKISISLPSMDL 408
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |

QY      225 DDLQETKIKINQLKIMTKGFTINSOPAYNGEKSDSPYGMGPGGYVQKAYVEFC 288
      1 | | | | : | | | : | | | | | | | | | | | | | | | | | | | |
DB      409 P-ITSEADILIRQLLISMKNNAFLTINSOPALNGEKSSHPFVGMPKPYFQKAYVEFV 467
      1 | | | | : | | | : | | | | | | | | | | | | | | | | | | | |

QY      289 SKKEDALVADKCDKRTSLTYMAVNMKDGSMKSNVGTQDVNATVYVFPKAKIIOPTIPDV 348
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |
DB      468 HRSLLNEKLEYTKKLNVSYFVNTKNGDLDINSQYEIRNATVWGVPFRREIIOPTIVEST 527
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |

QY      345 STNWKDFAFELWMSGKASLYPEDASRKLYVEVGGSHFLVSLVNDYVING-DLFAVF 405
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |
DB      528 SFLAKMDEYVSL-GMEAMNAYSPODISIRKLLVYIMNKKEMFLLVYIANDNDQNSQSLFDVF 584
      1 | | | | : | | | : | : | | | | | | | | : | | | | | | | | |

```

RESULT 4  
 S64136  
 methylmenetetrahydrofolate reductase (NADPH2) (EC 1.5.1.20) MET13 - yeast (*Saccharomyces*  
*Naltneri* name: protein G2882; protein YGL125W; ribosomal protein YML45, mitochond  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 03-Jun-2002  
 C:Accession: S64136; S53294; S78029; S72026  
 R:Cerdan, E.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Tizon, B.; Cadahia, J.L.  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S64134  
 A:Accession: S64136  
 A:Molecule type: DNA  
 A:Residues: 1-599 <CBER>  
 A:Cross-references: EMBL:D272647; NID:g1322686; PIDD:CA96833.1; PID:g1322687; MIPS:YG  
 A:Experimental source: strain S288C  
 R:Hosaka, K.; Nakawa, J.; Kodaki, T.; Ishizu, H.; Yamashita, S.  
 J. Biochem. 116, 1317-1321, 1994  
 A:Title: Cloning and sequence of the SCS3 gene which is required for inositol prototro  
 A:Reference number: JX0365; MUID:95221330; PMID:7706223  
 A:Accession: S53294  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-113, 'KRSS', <HOS>  
 A:Cross-references: EMBL:D12120  
 R:Itikawa, M.; Girack, H.R.; Grohmann, L.; Goldschmidt-Realsin, S.; Herfurth, E.; Wit  
 Eur. J. Biochem. 245, 449-456, 1997

A>Title: Identification and characterization of the genes for mitochondrial ribosomal p  
A:Reference number: S78018; MUID:97296414; PMID:9151978  
A:Accession: S78029  
A:Molecule type: protein  
A:Residues: 39-40; 'LA',43-47, 'A',49 <KIT>  
A>Note: this protein was identified as ribosomal protein Yml45, mitochondrial  
Ritzon, B.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Cadahia, J.L.; Cerdan, E.  
Yeast 12, 1047-1051, 1996  
A>Title: Identification of a putative methylentetrahydrofolate reductase by sequence an  
A:Reference number: S72026; MUID:97051592; PMID:8896269  
A:Accession: S72026  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-72; 'R',74-599 <TIZ>  
A:Cross-references: EMBL:X94106; NID:91628448; PIDN:CAA63833.1; PID:91628449  
A>Note: this protein was identified as putative methylentetrahydrofolate reductase  
C:Genetics:  
A:Gene: SGD:MET13; MET11; MRPL45  
A:Cross-references: MIPS:YGL125W; SGD:S0003093  
A:Map position: 7L  
A:Genome: nuclear  
C:Function: <METH>  
A:Description: oxidoreductase  
A>Note: this function seems to contradict the other function assigned to this protein  
C:Function: <RIB>  
A:Pathway: protein biosynthesis  
A:Note: this function seems to contradict the other function assigned to this protein  
C:Keywords: mitochondrion; oxidoreductase; protein biosynthesis; ribosome

Query Match 35.6%; Score 778; DB 2; Length 599;  
Best Local Similarity 40.9%; Pred. No. 3,5e-54;  
Matches 168; Conservative 76; Mismatches 147; Indels 20; Gaps 11;

QY 4 IYVQLFYDIDFLKRVNDCRIGITCPVPGIMNNYKGFIMTGFCKTKIPADIMAL 63  
DB 185 IITQMFYDIDFLKRVNDCRIGITCPVPGIMNNYKGFIMTGFCKTKIPADIMAL 244  
QY 64 EPIKDNKAEVAKAYGIHGLTGMCKKILAHG-IKTLHLTYLMKESALAILMNLGLI-EESK 121  
DB 245 DPKIDDELIVRIGITGNLIVEMCKKILDSGYSHLHYTMLEKAPMILIERILNLPTESE 304  
QY 122 VSRH-----LPMRRPANNFRKEDVRPIFMANRPKSYSTRIGW--DOYHGRMGDCSNPS 175  
DB 305 FNNHPLAVLPKRSKLNPRKNEEVRPIFMRKRPSTYAKRSQAVDEFPNGRGDSSSFA 364  
QY 176 YGAL-----SDYQMRPRARADKLLVEMAVPLKSYEDYERFRLYCLGKLSNPMWSELDL 231  
DB 365 FGDLDCGSD--LIRQSA--NKCLELWSTP-TSINDVAFVINYLNGNLKCLPMSDIP-I 418  
QY 232 QPEPTKILNQLEKINTKGFILINSOPAVNGEKSDSPVVGWGGGYYQKAYVEPSCSE 291  
DB 419 NDEINPIKRALHLELQHSIITINSOPVNGIRSKMDKIHGMPDGYVYQKHLEFRLPKT 478  
QY 292 KIDALVDKCKRSTSLTYMAVKNKDSWMSN-VGQTDNAVATMGVPAKEIIQPTIVDPVSF 350  
DB 479 KLPKRLDITKNNPFLTYFAIDSGDLISNHPDNKSNAAVWGIFFPREILOPFIYEKISF 538  
QY 351 NWWKDAEFIRSGWASLYPEDEA-SRKLYEVGSGFSLVLDNYINGD 400  
DB 539 LMKKEFFYIILNEMKLNMKYDKPHSAQFIQSLDDCYLVNIYDNDYISPD 589

RESULT 5  
T15423  
hypothetical protein C06A8.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T15423  
R:Leinbach, D.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid C06A8.  
A:Reference number: Z18348  
A:Accession: T15423

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-615 <LE>  
A:Cross-references: EMBL:U039849; NID:91055041; PID:91055045; PIDN:AAA81048.1; CESP:C0  
A:Genetics:  
A:Gene: CESP:C06A8.1  
A:Introns: 28/2; 53/2; 177/3; 375/3; 513/3; 565/2

Query Match 32.5%; Score 711; DB 2; Length 615;  
Best Local Similarity 38.4%; Pred. No. 8,3e-49;  
Matches 158; Conservative 63; Mismatches 126; Indels 64; Gaps 10;

QY 1 ARIYVQLFYDIDFLKRVNDCRIGITCPVPGIMNNYKGFIMTGFCKTKIPADIM 60  
DB 235 ANFVITQLFEAEFEPEKRVRCRREIGITQPLIPGIMPINGSYKRIALSQLLEIPQHL 294  
QY 61 AALEPKDNEEVAKAYGIHGLTGMCKKILAHG-IKTLHLTYLMKESALAILMNLGLIEE 119  
DB 295 DDLPIKDDDAVOKYGTERCIEKCRRLDNGTRAPSIHLTYMNRRESIREILKSLGL--- 351  
QY 120 SKVRSLSLPMRRPANNFRKEDVRPIFMANR--PKSYISRTIGMDQYPRGNGDCSNPSY 176  
DB 352 -----WK-----LEGDRVFPKMKNSQHPHOLBSTV----- 377  
QY 177 GALS DYQMRPRARADKLLVEMAVPLKSYEDYERFRLYCLG-----KLASNPMWSELD 229  
DB 378 -----RNADRLA-MEGANISFEDVAKRVPLINTYQAPNADGVKVTYLPTEAE 425  
QY 230 -GLQPEYKILNQLEKINTKGFILINSOPAVNGEKSDSPVVGWGGGYYQKAYVEEFC 288  
DB 426 TGVQPEYKILNQLEKINTKGFILINSOPAVNGEKSDSPVVGWGGGYYQKAYVEEFC 485  
QY 289 SKEKIDALVDKCKR--TSLTYMAVKNKDSM-KSNYGQTDNAVATMGVPAKEIIQPTIV 345  
DB 486 TAEISDKLQIIEKFEPRVRYHAIKNDSTFDKTNSETTIPAVITGVYPSGEIQAPIVY 545  
QY 346 DPVSFVNMKDEAEFIRSGWASLYPEDEASRKLVEEYGGSHFLSLVLDNDY 396  
DB 546 DPLSFRANWDEAYQWMAQMDGFYPKESKSGYVIAKAVHDFRLVTLVYDNDY 596

RESULT 6  
T38659  
methylentetrahydrofolate reductase 2 - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 20-Jun-2000  
C:Accession: T38659  
R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z21804  
A:Accession: T38659  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-641 <MDR>  
A:Cross-references: EMBL:AL109739; PIDN:CA82273.1; GSPDB:GN00066; SPDB:SPAC343.10  
A:Experimental source: strain 972h-; cosmid c343  
C:Genetics:  
A:Gene: SPDB:SPAC343.10  
A:Map position: 1

Query Match 28.8%; Score 629; DB 2; Length 641;  
Best Local Similarity 32.2%; Pred. No. 3,2e-42;  
Matches 147; Conservative 80; Mismatches 161; Indels 68; Gaps 13;

QY 1 ARIYVQLFYDIDFLKRVNDCRIGITCPVPGIMNNYKGFIMTGFCKTKIP 56  
DB 184 ADRITQIFERAEIKENFVRNHSNALRNPIIPALIMPQSYGSLKRMRLCGCVP 243  
QY 57 ADIMALEPKDNEEVAKAYGIHGLTGMCKKIL-----AHGKITLHLYTMKESALAI 110  
DB 244 SSLMORLNAKPPDEALKNIGVEHIVDKIKMDNVOGRVHG---FHFCTLIDERSVALI 300  
QY 111 LKNLGLI-----EESKVSRS-----LPMRRP-----ANVR 136

```

Db      301 LKNSGLITFRMKOYSEMEDEKLTTRKRLSLDEPAELHNOYVPSQGPVADKSSNLFV 360
OY      137 VAE-----DVPPIFWANRPKSYISRTIGMDQYPHGWMGSCSPGALSADY-- 182
Db      361 TSKOSSVSGHNDNLTEAP-FTVSESGVGLGRQANDDPFNGFDPRPAPGELDGTGP 419
OY      183 QPMRPARDKLYEEMAVPLKSVEDIYERRLYCLGLKNSNPWSELDGLQPTKILINOL 242
Db      420 TLHPFPEALKL--WGYPVDE-SDITSLFOKHIMSDISAPWID-EPYEVETKTIARYL 474
OY      243 EKINTKGLTINSOPAVNGEKSDSPYMGWGPAGYVOKAYVEFCCSKKELDALVDKCD 302
Db      475 LKNGSNMWTGSPAPVANGASADPVPFGWGPGRVQKAFVCEFNKDLDFITKMD 534
OY      303 RSTLYMAYNNKMGSKSVSGTVDNATWVPAKELIOPTIYDVPYFPMWDEAFETWS 362
Db      535 NQVITYYAGNNKSEFLTNAPKDGASATVWGVIYGRRELIGSTIIAEVSFKAMLSSEFYWG 594
OY      363 RGMASLYPEDEASKRLVEEVGSGHFLVSLVDNDYIN 398
Db      595 E-WANLYSKNTPSRKLENCINDRMVLYVIHDEMD 629

```

RESULT 7  
S63459  
methylene-tetrahydrofolate reductase (MADPH2) (EC 1.5.1.20) MET12 - yeast (Saccharomyces  
N/Alternate names: protein LPB8; protein YPL023c  
C/Species: Saccharomyces cerevisiae  
C/Date: 16-May-1996 #sequence\_revision 12-Jul-1996 #text\_change 03-Jun-2002  
C/Accession: S63459  
R/Wang, Y.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.; VC  
submitted to the EMBL Data Library, September 1995  
A/Reference number: S63452  
A/Accession: S63459  
A/Molecule type: DNA  
A/Residues: 1-657 <MAN>  
A/Cross-references: EMBL:U036624; NID:g1276642; PIDN:AAB68164.1; PID:g1039454; GSPDB:GN00  
C/Genetics:  
A/Genes: SGD:MET12; MIPS:YPL023c  
A/Cross-references: SGD:S0005944; MIPS:YPL023c  
A/Map position: 16L  
C/Keywords: oxidoreductase

```

Query Match      26.5%; Score 578.5; DB 2; Length 657;
Best Local Similarity 31.5%; Pred. No. 3.6e-38;
Matches 151; Conservative 80; Mismatches 155; Indels 93; Gaps 16;

OY      1 ARVITQLFYDTDFLKVNDRCQ-IGITCPYVGPINNYKGFTRMTGFCFKTIPADI 59
Db      183 ADFVITQLFYDEKFLTEMLFRRIISQDLPFPGIMPINSYLLFRAKLSHASIPAI 242
OY      60 MAALP-ISKDEAVKAYGILGTEMCKI---LHGIKTLHLTYLNMKSALATIM--- 112
Db      243 LSRFPEIGSDNNAKSGIVDILLELQEIYORTSGRIKGFHYTLNLEKAIQIVSGSP 302
OY      113 -----NLGLI-----EESKVSRSIPMRPANNV 134
Db      303 VLSHYNESSEEGDEFTSGEISIEVNPLEDADGDIVDDSSNEETVANRRK--RRHSL 360
OY      135 FRVKEDVREPIFWANR-----PKSYIS-----RTIGMDQYPHG 166
Db      361 ----SAAKLIF--NRAIVTEKGLRYNNENSGMPSKKALISIKGHGTLGRDATWDEFPNG 414
OY      167 RMGDSONSPYGLSDYQPMR--RARDKLYEEMAVPLKSVEDIYERFLYCLGLRBNP 224
Db      415 RFEDSRSAPYGEIDG---GPSIVSKSKALELMGIP-KTIGLKLIFIKYLEGSTDAP 470
OY      225 WSELGLOPETKTIINEOLEKINTKGLTINSOPAVNGEKSDSPYMGWGPAGYVOKAYV 284
Db      471 WSDL-GLSFTALIQEELIQLNRYGVLTLASQPATNATLSSDKIFGWMGPAKRLYOKAFV 529
OY      285 EFFCSKELDALVDKCD---RTSLTYMAYNNKMGSKSVSGTVDNATWVPAKELIO 341

```

```

Db      530 EMFIRQOEMTTLKPKLDHGRKRFSYAGDSSGSEFTMLDPHSSVYTWGVFPNRPVKQ 589
OY      342 PIIVPVSEFNWKKGEAFIWSRGMASLYPEDEASKRLVEEVGSGHFLVLDNDYIND 400
Db      590 TTIIEESFKARDEAFISME-WAKLFPRNTPANLLLVHKDCLVSYVHDFKED 647

```

RESULT 8  
S46454  
5,10-methylene-tetrahydrofolate reductase (FADH2) (EC 1.7.99.5) - human  
C/Species: Homo sapiens (man)  
C/Date: 15-Jul-1995 #sequence\_revision 17-Nov-1995 #text\_change 07-May-1999  
C/Accession: S46454  
R/Goyette, P.; Sumner, J.S.; Milos, R.; Duncan, A.M.V.; Rosenblatt, D.S.; Matthews, R  
Nature Genet. 7, 195-200, 1994  
A/Title: Human methylene-tetrahydrofolate reductase: isolation of cDNA, mapping and mu  
A/Reference number: S46454; M0ID:95004587; PMID:7920641  
A/Accession: S46454  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-416 <GOY>  
C/Keywords: FAD; oxidoreductase

```

Query Match      22.2%; Score 486; DB 2; Length 416;
Best Local Similarity 46.4%; Pred. No. 4.7e-31;
Matches 90; Conservative 35; Mismatches 67; Indels 2; Gaps 2;

OY      1 ARVITQLFYDTDFLKVNDRCQIGITCPYVGPINNYKGFTRMTGFCFKTIPADI 60
Db      223 ADFITQLFYDFEADVFRRVACADMGTCPYVGPISLQVLSKLELVEPEIK 282
OY      61 AALPEIKNEAVKAYGILGTEMCKILANG-IKTHLYTLNMKSALATIMGLTIE 119
Db      283 DYIEPIKNDAAINNYGIELAVSCQELASGLVPELHRYTLNRMATVEVLKRLGMWTE 342
OY      120 SKVSSLSLWRPANNVFKEDVPRIFWANRPKSYISRTIGMDQYPHGWMGSCSPGAL 179
Db      343 DP-RKPLWALSANPKREEDVPRIFWASPKSYITQEWDEFPNGMSSSPAREEL 401
OY      180 SDYQPMRPARDKR 193
Db      402 KDYYLFYLSKSPK 415

```

RESULT 9  
T42227  
methylene-tetrahydrofolate reductase homolog - fission yeast (Schizosaccharomyces pombe  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000  
C/Accession: T42227  
R/Yoshida, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.  
DNA Res. 4, 363-369, 1997  
A/Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.  
A/Reference number: 217323; M0ID:98162722; PMID:9501991  
A/Accession: T42227  
A/Status: preliminary; translated from CB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-566 <YOS>  
A/Cross-references: EMBL:D89118; NID:g1749443; PIDN:BA13780.1; PID:g1749444  
A/Experimental source: strain PR745

```

Query Match      19.3%; Score 421.5; DB 2; Length 566;
Best Local Similarity 30.6%; Pred. No. 1e-25;
Matches 110; Conservative 56; Mismatches 127; Indels 67; Gaps 12;

OY      1 ARVITQLFYDTDFLKVNDRCQIGITC---PIVGPIMPINNYKGFTRMTGFCFKTIP 56
Db      174 ADFITQLFYDFEAFIFEDNFVRNHSNALRNPIIPAIPIQSYGSLKRWTRLCGCSVP 233
OY      57 ADIMAALEPIDNEAVKAYGILGTEMCKIL-----AHGIKTLHLTYLNMKSALAI 110
Db      234 SSIQRLNAKRPDDEALKNIGVEHIYDMIKINDNQARVHG---FHCITLNRSLVALI 290

```

```

QY 111 LMLNLI-----ESKVSRS-----LPMRRP-----ANVR 136
D 291 LKMSGLITRKMKVSEKMEDEKLTTRKRLSDLEPAELHNOYVAVPQDPAKSSNLFV 350
QY 137 VKE-----DVREIFWANNRPKSYISRTIGMOQYPHGRKSGSCNPFYGLADY-- 182
D 351 TSKQSSVSGHKDNLFEAP-FVSESGSVLGRQANMDFNNGRRGDRSPAVGEIDGYCP 409
QY 183 QFMRRPRRDKLVEEAVNPLKSVEDITYERFLYLCIGLRSPMSSELDGLOPEFKIINED 242
D 410 TLHPPESEALKL---MGYPVDE-SDITSILFKHIMSDISVLPWHD-EPVEVEKRTIAKYL 464
QY 243 EKINTKQFLILINSOPVNGEKSPTSVMGGGGVYOKAVPEFCKSEKLDLALVYKCD 302
D 465 LKLGMSNMVTVGSOPVNGAPBAPDVFEGMGPKGGRVFOKPAVEFCVNGKILFFTKMHD 524

```

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.  
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak  
A.; Authors: da Silva, A.C.R.; da Silva F.R.; da Silva, A.M.; Silva Jr., M.A.; da Silva  
M.; Tshukko, M.H.; Valada, H.; van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A.; Reference number: A59328  
A; Contents: annotation  
C; Genetics:  
A; Gene: XF1121  
C; Superfamily: 5,10-methylenetetrahydrofolate reductase (FADH2)

Query Match 9.8%; Score 215; DB 2; Length 275;  
Best Local Similarity 33.9%; Pred. No. 1.2e-09;  
Matches 39; Conservative 30; Mismatches 46; Indels 0; Gaps 0;

Qy 1 ARVITQLEFYDIDFLKFNDCRQIGITCPYVGPIMPINNYKGFTRMTGCKTKIPADIM 60  
Db 160 ADAITQYFNPNFHFVDVAVRGVITPVGVPISNDKLNHFSQCGAELRWIT 219  
Qy 61 AALEPIKNEAVKAYGIGHLGTCKKILAHGIKTLHLTYLNMKSALAILMNLG 115  
Db 220 KKMAYGDDTKSTIRAFGADVATLCERLIAGAGLHFTYTLNLAKEPSTQVLRIG 274

## RESULT 14

5,10-methylenetetrahydrofolate reductase - Streptomyces coelicolor

C; Species: Streptomyces coelicolor  
C; Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jan-2000  
C; Accession: T34973  
R; Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1999  
A; Reference number: Z21563  
A; Accession: T34973  
A; Status: preliminary; translated from GB/EMBL/DBJ  
A; Molecule type: DNA  
A; Residues: 1-307 <SNV>  
A; Cross-references: EMBL:AL109663; PIDN:CAB52012.1; GSPDB:GN00070; SCOEDB:SC4A10.36C  
A; Experimental source: strain A3(2)  
C; Genetics:  
A; Gene: metF; SCOEDB:SC4A10.36C  
C; Superfamily: 5,10-methylenetetrahydrofolate reductase (FADH2)

Query Match 8.8%; Score 193; DB 2; Length 307;  
Best Local Similarity 33.0%; Pred. No. 8.3e-08;  
Matches 37; Conservative 22; Mismatches 53; Indels 0; Gaps 0;

Qy 5 VTQLFYDIDFLKFNDCRQIGITCPYVGPIMPINNYKGFTRMTGCKTKIPADIMALE 64  
Db 190 ITQMFQPDSDYRLRDVAAAGCATPVIPVPTSVYMLERLKLMSASFALKEKRL 249  
Qy 65 PIKNEAVKAYGIGHLGTCKKILAHGIKTLHLTYLNMKSALAILMNLG 116  
Db 250 TAKDDPAVRSIGIEFATFCARLLAEGVPGHFTLNNSTALEIENGL 301

## RESULT 15

5,10-methylenetetrahydrofolate reductase (FADH2) (BC 1.7.99.5) - Haemophilus influenzae

C; Species: Haemophilus influenzae  
C; Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C; Accession: H64123  
R; Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kervatage, A.  
; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Meldman, J.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
Science 269, 496-512, 1995  
A; Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A; Reference number: A64000; MUID: 95350630; PMID: 7542800  
A; Accession: H64123

A; Status: nucleic acid sequence not shown; translation not shown  
A; Molecule type: DNA  
A; Residues: 1-292 <TIGR>

C; Superfamily: 5,10-methylenetetrahydrofolate reductase (FADH2)  
C; Keywords: methionine biosynthesis; oxidoreductase

Query Match 8.6%; Score 189; DB 1; Length 292;  
Best Local Similarity 32.8%; Pred. No. 1.6e-07;  
Matches 38; Conservative 23; Mismatches 55; Indels 0; Gaps 0;

Qy 1 ARVITQLEFYDIDFLKFNDCRQIGITCPYVGPIMPINNYKGFTRMTGCKTKIPADIM 60  
Db 175 ANHVITQFFEDIENTLRFRDRCASIGIDTEIVGILPTNFKQKMSFTNVKIPAMLV 234  
Qy 61 AALEPIKNEAVKAYGIGHLGTCKKILAHGIKTLHLTYLNMKSALAILMNLG 116  
Db 235 KAYDGLNDPTTRNLVAASVAMDKILTSREGVNDPHTLRSELTVAICMILGV 290

Search completed: February 14, 2003, 23:10:48  
Job time : 44 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd

OM protein - protein search, using sw model

```
Run on:      February 14, 2003, 18:14:55 ; Search time 70 Seconds
              (without alignments)
              775.661 Million cell updates/sec
```

Title: US-09-720-451-6

Perfect score: 218/  
Sequence: 1 ARVIWQLFYDTDFLKEVN.....VSLVDNDYINGDLFAVFADF 408

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%

Database :

A\_Geneseq.101002.\*

1: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1980.DAT.\*

2: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1981.DAT.\*

3: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1982.DAT.\*

4: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1983.DAT.\*

5: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1984.DAT.\*

6: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1985.DAT.\*

7: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1986.DAT.\*

8: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1987.DAT.\*

9: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1988.DAT.\*

10: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1989.DAT.\*

11: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1990.DAT.\*

12: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1991.DAT.\*

13: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1992.DAT.\*

14: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1993.DAT.\*

15: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1994.DAT.\*

16: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1995.DAT.\*

17: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1996.DAT.\*

18: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1997.DAT.\*

19: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1998.DAT.\*

20: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA1999.DAT.\*

21: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA2000.DAT.\*

22: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA2001.DAT.\*

23: /SIDS2/gcgdata./geneseq./geneseq-emb1./AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2187	100.0	408	21	AAV44740	Soybean 5,10-methyl
2	1834		554	21	AAAG47726	Arabidopsis thaliana
3	1834	88.9	557	21	AAAG47725	Arabidopsis thaliana
4	1834	88.9	566	21	AAAG32205	Arabidopsis thaliana
5	1834	88.9	566	21	AAAG32204	Arabidopsis thaliana
6	1834	88.9	594	21	AAAG47724	Arabidopsis thaliana
7	1834	88.9	606	21	AAAG32203	Arabidopsis thaliana
8	1760	80.5	391	21	AAAG13450	Arabidopsis thaliana
9	1664	76.1	373	21	AAAG13451	Arabidopsis thaliana
10	1608	79.5	362	21	AAAG13452	Arabidopsis thaliana

11	1184	54.1	433	21	AAG36171	Arabidopsis thaliana
12	1184	54.1	470	21	AAG36170	Arabidopsis thaliana
13	1184	54.1	497	21	AAG36169	Arabidopsis thaliana
14	979	44.8	656	17	AAR88358	Human methylene-tetra
15	979	44.8	656	22	AAE12607	Human methylene-tetra
16	979	44.8	660	17	AAR88359	Human methylene-tetra
17	979	44.8	660	17	AAE16186	Human methylene-tetra
18	979	44.8	660	22	AAE12606	Human methylene-tetra
19	979	44.8	660	22	AAE15413	Human methylene-tetra
20	979	44.8	660	23	AAU75421	Human methylene-tetra
21	979	44.8	660	23	AAU75422	Human methylene-tetra
22	979	44.8	660	23	AAU75428	Human methylene-tetra
23	978	44.7	660	23	AAU75424	Human methylene-tetra
24	976	44.6	660	23	AAU75429	Human methylene-tetra
25	976	44.5	660	23	AAU75425	Human methylene-tetra
26	971	44.4	660	23	AAU75430	Human methylene-tetra
27	969	44.3	660	23	AAU75426	Human methylene-tetra
28	293	13.4	101	21	AA144741	Wheat 5.10-methylene
29	182	8.3	292	23	AAU72976	Nesleria meningitidis
30	163	7.5	283	23	ABBS4590	Lactococcus lactis
31	158	7.2	294	22	AAAG7970	Corynebacterium gl
32	158	7.2	326	22	AAAG92125	Corynebacterium prote
33	157	7.2	349	23	ABBO7695	C. glutamicum Meth
34	155.5	7.1	349	22	ABBS6361	Drosophila melanog
35	138	6.3	305	23	AAE13585	Streptomyces fradi
36	115	5.3	69	23	ABPO7534	Human ORFX prote
37	106.5	4.9	97	23	ABP11142	Human ORFX prote
38	103	4.7	1120	22	ABG16663	Novel human disgen
39	99.5	4.5	619	22	ABR64250	Vesicular stomatit
40	97.5	4.5	422	22	ABR20281	Drosophila melanog
41	97.5	4.5	2703	16	AAE70236	P. falciparum pro
42	97.5	4.5	2710	16	AAE22482	P. falciparum Pro3.
43	97.5	4.5	2710	21	AAV77904	P. falciparum Pro3.
44	97.5	4.5	3060	21	AAW22475	Plasmodium var-7.
45	97.5	4.5	3060	21	AAV77905	Plasmodium var-7

## ALIGNMENTS

## RESULT 1

ID	AA	Y44740	standard;	Protein;	408	AA.

AC AAY44740

DT 04-MAY-2000 (first entry)

DE Soybean 5,10-methylenetetrahydrofolate reductase.

KM 5,10-methylenetetrahydrofolate reductase; MTHFR; soybean; inhibitor  
KM methionine synthesis; tetrahydrofolate metabolism enzyme; herbicide

OS Glycine max.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT /note= "No codons in the corresponding nucleotide

Region	189..408
FT	

PN WO200004163-A1

PD 27-JAN-2000.

PF 14-JUL-1999; 99WO-US15916.

PR 15-JUL-1998; 98US-0092869.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Falco SC, Fomodu LO;  
XX  
DR WPI: 2000-182429/16.  
DR N-PSDB; AA250061.  
XX  
PI Novel tetrahydrofolate metabolism enzyme used to alter the level of  
PI tetrahydrofolate metabolism in plants and seeds -  
XX  
PS Claim 6; Pages 32-34; 37pp; English.  
XX  
CC The present sequence is a 5,10-methylenetetrahydrofolate reductase  
CC (MTHFR) encoded by cDNA clone sfil1.pk0017.d12. The clone was  
CC isolated from a cDNA library which was prepared using soybean  
CC immature flowers. MTHFR (E.C. 1.7.99.5) plays a role in the  
CC synthesis of methionine. The coding sequence of the enzyme is used in  
CC the construction of a chimeric gene to alter the level of  
CC tetrahydrofolate metabolism enzymes in plants. The enzyme may provide  
CC target to facilitate design and/or identification of inhibitors that may  
CC be useful as herbicides.  
CC N.B. The present sequence is also stated to be encoded by the  
CC ccol.pk0049.d4 clone in AA250060.  
XX  
SQ Sequence 408 AA;  
Query Match 100.0%; Score 2187; DB 21; Length 408;  
Best local similarity 100.0%; Pred. No. 1.2e-210;  
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAVITQLEFYDTDFLEKFNDCRQIGITCPIVPGIMPINNYKGFIRMTGCKTIPADIM 60  
DB 1 AAVITQLEFYDTDFLEKFNDCRQIGITCPIVPGIMPINNYKGFIRMTGCKTIPADIM 60  
QY 61 AALEPIKNEEYKAVYGIHGTGEMCKTILAHGKITLHLYTLNMEKSAIATMNGILEES 120  
DB 61 AALEPIKNEEYKAVYGIHGTGEMCKTILAHGKITLHLYTLNMEKSAIATMNGILEES 120  
QY 121 KVSRLPMRRPAPNFRKVEDRPTEFMANRKPSTISRTICMDQYPRGRWDCSNPSYGALS 180  
DB 121 KVSRLPMRRPAPNFRKVEDRPTEFMANRKPSTISRTICMDQYPRGRWDCSNPSYGALS 180  
QY 181 DVQFMRRAROKKLYEEMAVPLKSVEDIYERFRLYCLGRSLRNPSELDGLQPEKTIINE 240  
DB 181 DVQFMRRAROKKLYEEMAVPLKSVEDIYERFRLYCLGRSLRNPSELDGLQPEKTIINE 240  
QY 241 QLEKINTGFLTINSQPAVNGEKSDPTVGMGGPGGYVQKAYVEFFGSKKLDALYDKC 300  
DB 241 QLEKINTGFLTINSQPAVNGEKSDPTVGMGGPGGYVQKAYVEFFGSKKLDALYDKC 300  
QY 301 KQRTSLTYMAVVKDGSMSKSNVGGTDVNAVTVGVPAKETIIQPTIVDPVSFVWKDEAFET 360  
DB 301 KQRTSLTYMAVVKDGSMSKSNVGGTDVNAVTVGVPAKETIIQPTIVDPVSFVWKDEAFET 360  
QY 361 WSRGNASLYPEDEASRKLVEEVGSHFLVSLVNDYINGLDLEPAVADF 408  
DB 361 WSRGNASLYPEDEASRKLVEEVGSHFLVSLVNDYINGLDLEPAVADF 408  
RESULT 2  
AA647726  
ID AA647726 standard; Protein: 554 AA.  
XX  
AC AAG47726;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60184.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX

PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-012180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 08-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 21-APR-1999; 99US-0130077.  
PR 23-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135135.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0138119.  
PR 16-JUN-1999; 99US-0138452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.



```
RESULT 3
AG47725
ID AAC47725 standard; Protein: 557 AA.
XX
AC AAC47725;
XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60183.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hydriatation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PE
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 26-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148561.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
```

```
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161930.
PR 29-OCT-1999; 99US-0162142.

Query Match      83.9%; Score 1834; DB 21; Length 557;
Best Local Similarity 82.1%; Pred No. 57e-175;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;
```

```
QY 1 ARVYTOLEFDYDTFLKFNVDCKOIGTCPIVPGIMPINNKKGITRMGFCCKTKIPADIM 60
DB 150 ADLVYTOLEFYDTDFLKFVNDCKOIGTCPIVPGIMPINNKRGFLRMGFCCKTKIPVEYM 209
QY 61 AALEPIKDNBEAVKAYGIHGTBMCKKILAHGKITLHLYTLTLMESALATILMNLGLEES 120
DB 210 AALEPIKDNBEAVKAYGIHGTBMCKKMLAHGVKSLHLYTLTLMESALATILMNLGMIDES 269
QY 121 KYRSLSLPRRRRANFRYKKEVRPIFWANRPKYSISTRTIMWDYPIHGRNGDSCNFSYGALS 180
DB 270 KISRSLPWRPPANFRYKKEVRPIFWANRPKYSISTRTIMWDYPIHGRNGDSCNFSYGALS 329
```

```
QY 181 DYQMPRPARDKRLVEWMAVPLKSVEDIYERFRYLCJLKLKSNPMSELDTLOPETKINE 240
DB 330 DHQSRPRARDKRLQOEWVYPLKSVEDIOEKFKELCLGSLNKLSSPWSLDDLOPETKINE 389
QY 241 QLEKINTKGFLLTNSOPAVNGEKSDSPYWGCGPEGYIYOKAYIEFCSKEKLDALYDKC 300
DB 390 QLIKVNKSGFLTJNSQPSVNAERSDSPYWGCGPYGYOKAYIEFCSKEKLDAVEKC 449
QY 301 KDRTSLTYMAVNRKDSGSMKSNVCGQTDVNAVTVGWVPFAKEIIQPTIVDPVSFNWKADEAFET 360
DB 450 KALPSTIYMAVNRKGEQWVSNFTAQADVNAVTVGWVPFAKEIIQPTIVDPVSFNWKADEAFET 509
QY 361 WSRGMASTLYPEDASRKLYEEVGGSHFLSYLDNDYINGDIFANFAD 407
DB 510 WSRGMASTLYPEDASRRLLEEVKNSTLYLVENDYINGDIFANFAD 556

RESULT 4
AAG32205
ID AAG32205 standard; Protein; 566 AA.
AC AAG32205;
XX
XX 17-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38806.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PN
PD
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000BP-0301439.
PF
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
```

PR	27-MAY-1999	9905-01365992
PR	28-MAY-1999	9905-01367682
PR	01-JUN-1999	9905-01372222
PR	02-JUN-1999	9905-01375222
PR	04-JUN-1999	9905-01375528
PR	07-JUN-1999	9905-01377724
PR	08-JUN-1999	9905-01380094
PR	10-JUN-1999	9905-01385400
PR	10-JUN-1999	9905-01388477
PR	14-JUN-1999	9905-01391452
PR	15-JUN-1999	9905-01394652
PR	16-JUN-1999	9905-01394662
PR	18-JUN-1999	9905-01394662
PR	18-JUN-1999	9905-01394663
PR	18-JUN-1999	9905-01397560
PR	18-JUN-1999	9905-01397750
PR	21-JUN-1999	9905-01398177
PR	22-JUN-1999	9905-01398899
PR	23-JUN-1999	9905-01403553
PR	23-JUN-1999	9905-01403545
PR	24-JUN-1999	9905-01406655
PR	28-JUN-1999	9905-01408283
PR	29-JUN-1999	9905-01412897
PR	30-JUN-1999	9905-01412897
PR	01-JUL-1999	9905-01418482
PR	01-JUL-1999	9905-01421854
PR	02-JUL-1999	9905-01420354
PR	06-JUL-1999	9905-01422900
PR	08-JUL-1999	9905-01428083
PR	09-JUL-1999	9905-01429270
PR	12-JUL-1999	9905-01429700
PR	13-JUL-1999	9905-01435654
PR	14-JUL-1999	9905-01443642
PR	15-JUL-1999	9905-01444005
PR	15-JUL-1999	9905-01440895
PR	16-JUL-1999	9905-01440806
PR	13-JUL-1999	9905-01443315
PR	13-JUL-1999	9905-01443321
PR	19-JUL-1999	9905-01443432
PR	19-JUL-1999	9905-01443433
PR	19-JUL-1999	9905-01443434
PR	19-JUL-1999	9905-01443435
PR	19-JUL-1999	9905-01443435
PR	20-JUL-1999	9905-01443532
PR	20-JUL-1999	9905-01446532
PR	20-JUL-1999	9905-01448884
PR	21-JUL-1999	9905-01448814
PR	21-JUL-1999	9905-01450806
PR	21-JUL-1999	9905-01450808
PR	21-JUL-1999	9905-01450885
PR	22-JUL-1999	9905-01450889
PR	22-JUL-1999	9905-01450897
PR	22-JUL-1999	9905-01451545
PR	22-JUL-1999	9905-01451542
PR	23-JUL-1999	9905-01452224
PR	23-JUL-1999	9905-01452228
PR	26-JUL-1999	9905-01455776
PR	27-JUL-1999	9905-01455913
PR	27-JUL-1999	9905-01455918
PR	27-JUL-1999	9905-01455919
PR	28-JUL-1999	9905-01455919
PR	02-AUG-1999	9905-01446588
PR	02-AUG-1999	9905-01446588
PR	02-AUG-1999	9905-01446588
PR	03-AUG-1999	9905-01447038

PR	04-ANG-1999	9905-0147204
PR	04-ANG-1999	9905-0147302
PR	05-ANG-1999	9905-0147192
PR	05-ANG-1999	9905-0147760
PR	06-ANG-1999	9905-0147730
PR	06-ANG-1999	9905-0147436
PR	09-ANG-1999	9905-0147619
PR	09-ANG-1999	9905-0147735
PR	10-ANG-1999	9905-0148171
PR	11-ANG-1999	9905-0148341
PR	12-ANG-1999	9905-0148365
PR	13-ANG-1999	9905-0148684
PR	16-ANG-1999	9905-0149368
PR	17-ANG-1999	9905-0149175
PR	18-ANG-1999	9905-0149426
PR	20-ANG-1999	9905-0149722
PR	20-ANG-1999	9905-0149723
PR	22-ANG-1999	9905-0149902
PR	23-ANG-1999	9905-0149930
PR	25-ANG-1999	9905-0150566
PR	26-ANG-1999	9905-0150884
PR	27-ANG-1999	9905-0151065
PR	27-ANG-1999	9905-0151067
PR	27-ANG-1999	9905-0151080
PR	30-ANG-1999	9905-0151303
PR	31-SEP-1999	9905-0151338
PR	01-SEP-1999	9905-0151360
PR	07-SEP-1999	9905-0152673
PR	10-SEP-1999	9905-0153070
PR	13-SEP-1999	9905-0153718
PR	15-SEP-1999	9905-0154018
PR	16-SEP-1999	9905-0154039
PR	20-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155139
PR	23-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0155539
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156568
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157713
PR	06-OCT-1999	9905-0157655
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158323
PR	12-OCT-1999	9905-0158569
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159331
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	16-OCT-1999	9905-0159584
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160767
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160770
PR	21-OCT-1999	9905-0160814
PR	21-OCT-1999	9905-0160815
PR	22-OCT-1999	9905-0160880
PR	22-OCT-1999	9905-0160981
PR	22-OCT-1999	9905-0160982
PR	25-OCT-1999	9905-0161404
PR	25-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	26-OCT-1999	9905-0161359
PR	26-OCT-1999	9905-0161360
PR	26-OCT-1999	9905-0161361
PR	28-OCT-1999	9905-0161992
PR	28-OCT-1999	9905-0161993



PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145089
PR	22-JUL-1999	9905-0145152
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145524
PR	26-JUL-1999	9905-0145576
PR	27-JUL-1999	9905-0145513
PR	27-JUL-1999	9905-0145518
PR	28-JUL-1999	9905-0145591
PR	02-AUG-1999	9905-0146386
PR	02-AUG-1999	9905-0146388
PR	02-AUG-1999	9905-0146389
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147435
PR	09-AUG-1999	9905-0147593
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148319
PR	13-AUG-1999	9905-0148564
PR	13-AUG-1999	9905-0148684
PR	16-AUG-1999	9905-0149376
PR	17-AUG-1999	9905-0149191
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149929
PR	22-AUG-1999	9905-0149932
PR	23-AUG-1999	9905-0149930
PR	23-AUG-1999	9905-0149930
PR	26-AUG-1999	9905-0150566
PR	27-AUG-1999	9905-0150884
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151338
PR	01-SEP-1999	9905-0151340
PR	07-SEP-1999	9905-0152363
PR	10-SEP-1999	9905-0153070
PR	13-SEP-1999	9905-0153158
PR	15-SEP-1999	9905-0154079
PR	16-SEP-1999	9905-0154739
PR	20-SEP-1999	9905-0155139
PR	22-SEP-1999	9905-0155186
PR	24-SEP-1999	9905-0155659
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157553
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158332
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	14-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	18-OCT-1999	9905-0159584
PR	21-OCT-1999	9905-0160741

PR	21-OCT-1999	99US-0160767
PR	21-OCT-1999	99US-0160768
PR	21-OCT-1999	99US-0160770
PR	21-OCT-1999	99US-0160814
PR	21-OCT-1999	99US-0160815
PR	22-OCT-1999	99US-0160980
PR	22-OCT-1999	99US-0160981
PR	22-OCT-1999	99US-0160982
PR	22-OCT-1999	99US-0160984
PR	25-OCT-1999	99US-0161404
PR	25-OCT-1999	99US-0161405
PR	25-OCT-1999	99US-0161406
PR	25-OCT-1999	99US-0161407
PR	26-OCT-1999	99US-0161350
PR	26-OCT-1999	99US-0161361
PR	28-OCT-1999	99US-0161320
PR	28-OCT-1999	99US-0161322
PR	28-OCT-1999	99US-0161343
PR	29-OCT-1999	99US-0162142

Query Match	83.9%;	Score 1834;	DB 21;	Length 569;
Best Local Similarity	82.1%;	Pred. No. 5.9e-175;		
Matches 334;	Conservative 37;	Mismatches 36;	Indels 0;	Gaps 0.

[illegible]

XX	RESULT 6
XX	AAAG47724
ID	AAAG47724 standard; Protein: 594 AA.
XX	
AC	AAAG47724:
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 60182.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
EN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.

XX	25-FEB-1999.	99US-0121825.	PR	12-JUL-1999.	99US-0142977.
PR	05-MAR-1999.	99US-0123180.	PR	13-JUL-1999.	99US-0143342.
PR	03-MAR-1999.	99US-0123548.	PR	14-JUL-1999.	99US-0143624.
PR	23-MAR-1999.	99US-0125788.	PR	15-JUL-1999.	99US-0144005.
PR	25-MAR-1999.	99US-0126264.	PR	16-JUL-1999.	99US-0144085.
PR	29-MAR-1999.	99US-0126785.	PR	17-JUL-1999.	99US-0144086.
PR	01-APR-1999.	99US-0127462.	PR	19-JUL-1999.	99US-0144325.
PR	06-APR-1999.	99US-0128234.	PR	19-JUL-1999.	99US-0144331.
PR	08-APR-1999.	99US-0128714.	PR	19-JUL-1999.	99US-0144332.
PR	16-APR-1999.	99US-0129845.	PR	19-JUL-1999.	99US-0144333.
PR	19-APR-1999.	99US-0130077.	PR	19-JUL-1999.	99US-0144334.
PR	21-APR-1999.	99US-0130449.	PR	19-JUL-1999.	99US-0144335.
PR	23-APR-1999.	99US-0130510.	PR	20-JUL-1999.	99US-0144352.
PR	28-APR-1999.	99US-0130891.	PR	20-JUL-1999.	99US-0144632.
PR	30-APR-1999.	99US-0131449.	PR	20-JUL-1999.	99US-0144684.
PR	30-APR-1999.	99US-0132048.	PR	21-JUL-1999.	99US-0144814.
PR	04-MAY-1999.	99US-0132407.	PR	21-JUL-1999.	99US-0145086.
PR	05-MAY-1999.	99US-0132484.	PR	22-JUL-1999.	99US-0145088.
PR	06-MAY-1999.	99US-0132485.	PR	22-JUL-1999.	99US-0145085.
PR	06-MAY-1999.	99US-0132486.	PR	22-JUL-1999.	99US-0145087.
PR	07-MAY-1999.	99US-0132487.	PR	22-JUL-1999.	99US-0145089.
PR	11-MAY-1999.	99US-0132863.	PR	23-JUL-1999.	99US-0145145.
PR	11-MAY-1999.	99US-0134256.	PR	23-JUL-1999.	99US-0145145.
PR	14-MAY-1999.	99US-0134218.	PR	23-JUL-1999.	99US-0145218.
PR	14-MAY-1999.	99US-0134219.	PR	26-JUL-1999.	99US-0145324.
PR	14-MAY-1999.	99US-0134221.	PR	26-JUL-1999.	99US-0145376.
PR	14-MAY-1999.	99US-0134370.	PR	27-JUL-1999.	99US-0145513.
PR	18-MAY-1999.	99US-0134768.	PR	27-JUL-1999.	99US-0145518.
PR	18-MAY-1999.	99US-0134941.	PR	27-JUL-1999.	99US-0145519.
PR	20-MAY-1999.	99US-0135124.	PR	28-JUL-1999.	99US-0145591.
PR	21-MAY-1999.	99US-0135353.	PR	02-AUG-1999.	99US-0146386.
PR	24-MAY-1999.	99US-0135629.	PR	02-AUG-1999.	99US-0146388.
PR	25-MAY-1999.	99US-0136021.	PR	02-AUG-1999.	99US-0146389.
PR	27-MAY-1999.	99US-0136392.	PR	03-AUG-1999.	99US-0147038.
PR	28-MAY-1999.	99US-0136782.	PR	04-AUG-1999.	99US-0147204.
PR	01-JUN-1999.	99US-0137222.	PR	04-AUG-1999.	99US-0147302.
PR	03-JUN-1999.	99US-0137502.	PR	05-AUG-1999.	99US-0147192.
PR	04-JUN-1999.	99US-0137502.	PR	05-AUG-1999.	99US-0147260.
PR	07-JUN-1999.	99US-0137724.	PR	06-AUG-1999.	99US-0147303.
PR	08-JUN-1999.	99US-0138094.	PR	06-AUG-1999.	99US-0147416.
PR	10-JUN-1999.	99US-0138540.	PR	09-AUG-1999.	99US-0147493.
PR	10-JUN-1999.	99US-0138847.	PR	09-AUG-1999.	99US-0147935.
PR	14-JUN-1999.	99US-0139119.	PR	10-AUG-1999.	99US-0148171.
PR	16-JUN-1999.	99US-0139452.	PR	11-AUG-1999.	99US-0148319.
PR	16-JUN-1999.	99US-0139453.	PR	12-AUG-1999.	99US-0148341.
PR	17-JUN-1999.	99US-0139452.	PR	13-AUG-1999.	99US-0148365.
PR	18-JUN-1999.	99US-0139454.	PR	13-AUG-1999.	99US-0148684.
PR	18-JUN-1999.	99US-0139455.	PR	16-AUG-1999.	99US-0149368.
PR	18-JUN-1999.	99US-0139456.	PR	17-AUG-1999.	99US-0149175.
PR	18-JUN-1999.	99US-0139457.	PR	18-AUG-1999.	99US-0149426.
PR	18-JUN-1999.	99US-0139458.	PR	20-AUG-1999.	99US-0149722.
PR	18-JUN-1999.	99US-0139459.	PR	20-AUG-1999.	99US-0149723.
PR	18-JUN-1999.	99US-0139460.	PR	20-AUG-1999.	99US-0149829.
PR	18-JUN-1999.	99US-0139461.	PR	23-AUG-1999.</	

PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159293.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159684.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160766.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161933.  
PR 29-OCT-1999; 99US-0162142.

Query Match 83.9%; Score 1834; DB 21; Length 594;  
Best Local Similarity 82.1%; Pred. No. 6,3e-175;  
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

QY 1 ARVIYTOLEFYDDIFLKFVNDRCQIGITGPYIPGIMPINNYGFLRMGFECKTKIPADIM 60  
DB 187 ADLIVTOLFEDYDIFLKFVNDRCQIGISCPYIPGIMPINNYGFLRMGFECKTKIPADIM 246  
QY 61 AALEPIDNEEAVKAGIHGTEMCKKILAHGKTLHYTLNMEKSALAILNLGLIEES 120  
DB 247 AALEPIDNEEAVKAGIHGTEMCKKILAHGKSHLYTLNMEKSALAILNLGLIMDES 306  
QY 121 KVSRLPWRPRAVFRYKEDVRPIFWANRPKSYISTIGMDQYPHGRMGDSGNPSYGALS 180  
DB 307 KISRLPWRPRAVFRYKEDVRPIFWANRPKSYISTRTKGMEDFPOGRMGDSRSASYGALS 366  
QY 181 DYQMRPARRADKILVEEMAVPLKSVEDIERFRLYCLGKLRSPMSELDGLQPEKRIINE 240  
DB 367 DHQSRPARRADKILQOQWVPLKSVEDIQEKRELCLGNKSSPMSELDGLQPEKRIINE 426  
QY 241 QLEKINTKGFLLINSQPAVNGEKSDEPTVWGPGGYVYQKAYVEFFCSKEKLDALVDKC 300  
DB 427 QLIIVNSKGFLLINSQPSVAHERSDSPTVWGPGGYVYQKAYLEFFCSKEKLDANAVEKC 486  
QY 301 KDRTSLTYMAVNVKSGSKSVNGOTDVNAVWGFPAKEITIOPTIVPVFNFNWKDEAFPI 360  
DB 487 KALPSITYMAVNVKSGEQVNVSTADVNAVWGFPAKEITIOPTIVPVFNFNWKDEAFET 546  
QY 361 WSRGMASLYPEDEASRKLVFEGVSHFLVSLVNDNDYINGDLFAVFAD 407  
DB 547 WSRGMASLYPEADPSRNLEEVKNSYTLVSLVENDYINGDIFAVFAD 593

RESULT 7  
AAG32203  
ID AAG32203 standard; Protein; 606 AA.  
XX

AC AAG32203;  
XX 17-OCT-2000 (first entry)  
DT Arabidopsis thaliana protein fragment SEQ ID NO: 38804.  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38804.  
DE Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
PR 04-MAY-1999; 99US-0132407.  
XX 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132486.  
XX 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
XX 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
XX 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
XX 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
XX 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
XX 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
XX 18-JUN-1999; 99US-0139462.





PR	06-AUG-1999;	9905-0147416.
PR	09-AUG-1999;	9905-0147493.
PR	09-AUG-1999;	9905-0147935.
PR	10-AUG-1999;	9905-0148171.
PR	11-AUG-1999;	9905-0148319.
PR	12-AUG-1999;	9905-0148341.
PR	13-AUG-1999;	9905-0148565.
PR	13-AUG-1999;	9905-0148684.
PR	16-AUG-1999;	9905-0149368.
PR	17-AUG-1999;	9905-0149175.
PR	18-AUG-1999;	9905-0149426.
PR	20-AUG-1999;	9905-0149722.
PR	20-AUG-1999;	9905-0149723.
PR	20-AUG-1999;	9905-0149929.
PR	23-AUG-1999;	9905-0149902.
PR	23-AUG-1999;	9905-0149930.
PR	25-AUG-1999;	9905-0150566.
PR	26-AUG-1999;	9905-0150884.
PR	27-AUG-1999;	9905-0151065.
PR	27-AUG-1999;	9905-0151066.
PR	30-AUG-1999;	9905-0151080.
PR	30-AUG-1999;	9905-0151303.
PR	31-AUG-1999;	9905-0151438.
PR	01-SEP-1999;	9905-0151930.
PR	07-SEP-1999;	9905-0152363.
PR	10-SEP-1999;	9905-0153070.
PR	13-SEP-1999;	9905-0153758.
PR	15-SEP-1999;	9905-0154018.
PR	16-SEP-1999;	9905-0154039.
PR	20-SEP-1999;	9905-0154779.
PR	22-SEP-1999;	9905-0155139.
PR	23-SEP-1999;	9905-0155486.
PR	24-SEP-1999;	9905-0155659.
PR	28-SEP-1999;	9905-0156458.
PR	29-SEP-1999;	9905-0156596.
PR	04-OCT-1999;	9905-0157117.
PR	05-OCT-1999;	9905-0157753.
PR	06-OCT-1999;	9905-0157865.
PR	07-OCT-1999;	9905-0158029.
PR	08-OCT-1999;	9905-0158332.
PR	12-OCT-1999;	9905-0158369.
PR	13-OCT-1999;	9905-0159293.
PR	13-OCT-1999;	9905-0159294.
PR	13-OCT-1999;	9905-0159295.
PR	14-OCT-1999;	9905-0159329.
PR	14-OCT-1999;	9905-0159330.
PR	14-OCT-1999;	9905-0159331.
PR	14-OCT-1999;	9905-0159637.
PR	14-OCT-1999;	9905-0159638.
PR	18-OCT-1999;	9905-0159584.
PR	21-OCT-1999;	9905-0160741.
PR	21-OCT-1999;	9905-0160767.
PR	21-OCT-1999;	9905-0160768.
PR	21-OCT-1999;	9905-0160770.
PR	21-OCT-1999;	9905-0160814.
PR	21-OCT-1999;	9905-0160815.
PR	22-OCT-1999;	9905-0160980.
PR	22-OCT-1999;	9905-0160981.
PR	22-OCT-1999;	9905-0160989.
PR	23-OCT-1999;	9905-0161404.
PR	25-OCT-1999;	9905-0161405.
PR	25-OCT-1999;	9905-0161406.
PR	26-OCT-1999;	9905-0161359.
PR	26-OCT-1999;	9905-0161360.
PR	26-OCT-1999;	9905-0161361.
PR	28-OCT-1999;	9905-0161920.
PR	28-OCT-1999;	9905-0161920.
PR	28-OCT-1999;	9905-0161993.
PR	29-OCT-1999;	9905-0162142.

QY		18	FVNDCRQIGITPIYPGIMPINNYGFIIRMGFCCKTIKIPADIMALEPIKDNEAVAYG	77
Db		1	VFNDRRGIGINCPYIPGIMPIPNNYRGFLRMFGFCKTKIPEVMMALEPIKDNEEVAAYG	60
QY		78	IHLGTGEMCKKLIAHGIKTILHYTLTUMMEKSALATLMNLGLIEESKYSRSLPMPRRANPRV	137
Db		61	IHLGTGEMCKKLIAHGKVSLSLHTLTUMMEKSALATLMNLGMIDESKISRSLPMRRANVFRT	120
QY		138	KEDVPPIEFMANPKSYISRTIKGMDQYPHGRMGDSQCNPSYGALSDPYOFMRPARDKIYEE	197
Db		121	KEDVAPIEFMANPKSYISRTKGMEDEFPQGRMGDSASAYGALSDBHQSRPARAPKKIQOE	180
QY		198	NAPVLKSEVDIYEERFLYLCLGRSNPWSSELDGLOPETKIINEOLEKINTKGFLTINSQP	257
Db		181	NWVPLKSEVDIEDKEKEKELCTLGMLKSSPMSSELGDLOPETKIINEOLYVKNSGFLTINSQP	240
QY		258	AVNGEKSSSPYVGWGPGGYGYQKAYVEFPFSKEKLDLVCKCDRSLLTYMANKGSW	317
Db		241	SVNAERSDSPYVGWGPVGYGYQKAYLEFFCSKEKLDLVVERCKALPSITYMAVNKGEOW	300
QY		318	KSNVQTQDVNATVMTGVFPFAKETIOPTIYDPSFNWVKDEAFETISRCMASLYPEDEASRK	377
Db		301	VSNVQAQADVNAVNTVGVFPFAKETIOPTIYDPASFNVWVKDEAFETISRCMANLYPEADSRN	360
QY		378	LVEEVGGSHFLVSLVDNDYINGDLFAVPAAD	407
Db		361	LLEEVKNSTYLVSLVENDYINGDLFAVPAAD	390
 RESULT 9 AAG15451				
ID	AAG15451	standard; Protein;	373 AA.	
XX				
AC	AAG15451;			
XX				
DT	17-OCT-2000	(first entry)		
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 15708.		
XX				
KW	Protein identification; signal transduction pathway; metabolic pathway;			
KM	hybridisation assay; genetic mapping; gene expression control; promoter;			
KW	termination sequence.			
XX				
OS	Arabidopsis thaliana.			
PN	EP1033405-A2.			
XX				
PD	06-SEP-2000.			
XX				
PF	25-FEB-2000; 2000EP-0301439.			
XX				
FR	25-FEB-1999;	99US-0121825.		
FR	05-MAR-1999;	99US-0123180.		
FR	09-MAR-1999;	99US-0123548.		
PR	23-MAR-1999;	99US-0125788.		
PR	25-MAR-1999;	99US-0126284.		
PR	29-MAR-1999;	99US-0126785.		
PR	01-APR-1999;	99US-0127462.		
PR	06-APR-1999;	99US-0128234.		
PR	08-APR-1999;	99US-0128714.		
PR	16-APR-1999;	99US-0129884.		
PR	19-APR-1999;	99US-0130077.		
PR	21-APR-1999;	99US-0130449.		
PR	23-APR-1999;	99US-0130510.		
PR	23-APR-1999;	99US-0130891.		
PR	28-APR-1999;	99US-0131449.		
PR	30-APR-1999;	99US-0132048.		
PR	30-APR-1999;	99US-0132407.		
PR	04-MAY-1999;	99US-0132484.		
PR	05-MAY-1999;	99US-0132485.		
PR	06-MAY-1999;	99US-0132486.		
PR	06-MAY-1999;	99US-0132487.		

PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144614.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 76.1%; Score 1664; DB 21; Length 373;  
Best Local Similarity 81.5%; Pred. No. 3 6e-158;  
Matches 303; Conservative 34; Mismatches 35; Indels 0; Gaps 0;

OY 36 MPNNKGFIRMTGFCCKTIPADIMALEPIKDNBEAVKAYGIHLGTEMCKKILAHGIKT 95  
DB 1 MPNNKGFIRMTGFCCKTIPYEMMALEPIKDNBEAVKAYGIHLGTEMCKKILAHGVKS 60  
OY 96 LHIYTNMKSALAILMNLGLIEESKVSRLPWRPRAVFERKEDVRIFFMANRKSYS 155  
DB 61 LHIYTNMKSALAILMNLGLIEESKVSRLPWRPRAVFERKEDVRIFFMANRKSYS 120  
OY 156 RTIGMDYPHGRWGDSCNPYCALSDYOPRRPADKLYEEMAPLKSVEDIYERFRLY 215  
DB 121 RTIGWEDFPOGRWGDSCNPYCALSDHOSRRPADKLYEEMAPLKSVEDIYERFRLY 180  
OY 216 CIGKLRSNPWSFLDGLQPTKILINOLEKINTKGLTINSOPAVNGEKSDSTVWGSGPG 275  
DB 181 CIGKLRSNPWSFLDGLQPTKILINOLEKINTKGLTINSOPAVNGEKSDSTVWGSGPG 240  
OY 276 GYVYOKAYVEFCSEKELDALYDKCKDRTSLTYMAVNNKGSMKSNVGGTDVNAVTVGWGEP 335  
DB 241 GYVYOKAYVEFCSEKELDALYDKCKDRTSLTYMAVNNKGSMKSNVGGTDVNAVTVGWGEP 300  
OY 336 AKETIIOPTIVDPVSFNWMDKDEAFELIWSRGWASLYPEDASRKLVEVGSHFLVSLVND 395  
DB 301 AKETIIOPTIVDPVSFNWMDKDEAFELIWSRGWASLYPEDASRKLVEVGSHFLVSLVND 360  
OY 396 YINGDLFAVFAAD 407  
DB 361 YINGDLFAVFAAD 372

RESULT 10  
AAG15452  
ID AAG15452 standard; Protein: 362 AA.

AC AAG15452;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 15709.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123160.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130049.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140921.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.

PR	15-JUL-1999;	9905-0144086;
PR	19-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144352;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144864;
PR	21-JUL-1999;	9905-0144814;
PR	21-JUL-1999;	9905-0145086;
PR	22-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145593;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	28-JUL-1999;	9905-0145961;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148177;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	26-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151530;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155139;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155659;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156516;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158029;

PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161982.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	73.5%	Score 1608	DB 21	Length 362
Best Local Similarity	81.4%	Pred. No. 1,4e-152		
Matches 294	Conservative 32	Mismatches 35	Indels 0	Gaps 0
QY	47	MTGFCKTKIPADIMAALEPIKDNEEAVKAYGIHLGTEMCKKIIANGIKTLHLTYLNMKS	106	
DB	1	MTGFCKTKIPYEMAALEPIKDNEEAVKAYGIHLGTEMCKKMAHGVKSLHLTYLNMKS	60	
QY	107	ALATIMNIGLIEESVNSLSLPMRRPRAVFEKEDVPIPIANRPKSYISRTIGMOPYHG	166	
DB	61	ALATIMNIGMIDESKISLSLPMRRPRAVFEKEDVPIPIANRPKSYISRTIGMEDPFG	120	
QY	167	RMGDSGNSPYASLDYQFMRRARDKKLVEEMAVPLKSVEDIYERERLYLGLKLSNPM	226	
DB	121	RMGDSRSASYGALSDDHGRSRPRARDKKLQDEWVYPLKSVEDIQEFKELCLNTLKSP	180	
QY	227	ELDDGIQPEPKTIINDELEKINTKGFLLTINSQPAVNGEKSDSPYGMGPGGAYOKRAYE	286	
DB	181	ELDDGIQPEPKTIINDELYVNVSKGFLTINSQSVAAEKSDSPYGMGPGGAYOKRAYE	240	
QY	287	PCSKEKLDALVDCCKDRFSLTYMAVNVKDGSMKSNVGGTQDVNAVATWGVFPFAKEIIQ	346	
DB	241	PCSKEKLDAYVEKCKALPSIITYMAVNVNGEQWVSNTYADVNAVATWGVFPFAKEIIQ	300	
QY	347	PVSFNWMDAEAFELMSRGMASLYPEDDASKRLVDEVGSGFLVSYLDNDYINGDILFAVA	406	
DB	301	PASFNWMDAEAFETMSRMANLYPEDDPSRLLLEEVANSYLVLSVLENDYINGDILFAVA	360	
QY	407	P 407		
DB	361	P 361		
RESULT 11				
AAAG36171				
ID	AAAG36171	standard; Protein; 433 AA.		
XX	AAAG36171;			
AC				
XX				
DT	18-OCT-2000	(first entry)		
XX				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 44290.			

XX Protein identification: signal transduction pathway; metabolic pathway;  
KM hybridisation assay: genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 990S-0121825.  
PR 05-MAR-1999; 990S-01233180.  
PR 09-MAR-1999; 990S-0123548.  
PR 23-MAR-1999; 990S-0125788.  
PR 25-MAR-1999; 990S-0126264.  
PR 29-MAR-1999; 990S-0126785.  
PR 01-APR-1999; 990S-0127462.  
PR 06-APR-1999; 990S-0128234.  
PR 08-APR-1999; 990S-0128714.  
PR 16-APR-1999; 990S-0129845.  
PR 19-APR-1999; 990S-0130077.  
PR 21-APR-1999; 990S-01300449.  
PR 23-APR-1999; 990S-0130510.  
PR 23-APR-1999; 990S-0130891.  
PR 28-APR-1999; 990S-0131449.  
PR 30-APR-1999; 990S-0132048.  
PR 30-APR-1999; 990S-0132407.  
PR 04-MAY-1999; 990S-0132484.  
PR 05-MAY-1999; 990S-0132485.  
PR 06-MAY-1999; 990S-0132486.  
PR 07-MAY-1999; 990S-0132487.  
PR 11-MAY-1999; 990S-0132863.  
PR 14-MAY-1999; 990S-0134256.  
PR 14-MAY-1999; 990S-0134218.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136392.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137232.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137724.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.  
PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 16-JUN-1999; 990S-0139453.  
PR 17-JUN-1999; 990S-0139452.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139763.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.

PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 29-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 01-JUL-1999; 990S-0142154.  
PR 02-JUL-1999; 990S-0142055.  
PR 06-JUL-1999; 990S-0142380.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142920.  
PR 12-JUL-1999; 990S-0142977.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 20-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 20-JUL-1999; 990S-0144884.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145224.  
PR 26-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 02-AUG-1999; 990S-0146389.  
PR 03-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 04-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 05-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 06-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147493.  
PR 09-AUG-1999; 990S-0147935.  
PR 10-AUG-1999; 990S-0148171.  
PR 11-AUG-1999; 990S-0148319.  
PR 12-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 13-AUG-1999; 990S-0148684.  
PR 16-AUG-1999; 990S-0149368.  
PR 17-AUG-1999; 990S-0149175.  
PR 18-AUG-1999; 990S-0149426.  
PR 20-AUG-1999; 990S-0149722.  
PR 20-AUG-1999; 990S-0149723.  
PR 20-AUG-1999; 990S-0149929.  
PR 23-AUG-1999; 990S-0149902.  
PR 23-AUG-1999; 990S-0149930.  
PR 25-AUG-1999; 990S-0150566.  
PR 26-AUG-1999; 990S-0150884.  
PR 27-AUG-1999; 990S-0151065.  
PR 27-AUG-1999; 990S-0151066.  
PR 27-AUG-1999; 990S-0151080.  
PR 30-AUG-1999; 990S-0151303.

PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159337.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 25-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 54.1%; Score 1184; DB 21; Length 433;  
Best Local Similarity 78.9%; Pred. No. 7.7e-110;

Matches 213; Conservative 28; Mismatches 29; Indels 0; Gaps 0;

QY 1 ARVITQLFYDIDIFLKFVNDCRQIGITCPVPGIMPINNYKGFIRMTGFCKTIPADIM 60  
|:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|  
Db 150 ADIVTQLFYDIDIFLKFVNDCRQIGITCPVPGIMPISNKGFLRMAGFCKTIPAELE 209  
QY 61 AALEPTIDNEEAVKAGCIGILGTECKKILAHGKIKTLHLYTNMESALAIIMNGLJES 120  
|:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|  
Db 210 AALEPTIDNEEAVKAGCIGILGTECKKILAHGKIKTLHLYTNMESALAIIMNGLJES 269  
QY 121 KVSLSLWRRPAPVFEVKEDVRPIFWANRPKSYISRTIGMDQYPHRGWGSCLNPSTGALS 180  
|:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|  
Db 270 KISLSLWRRPAPVFEVKEDVRPIFWANRPKSYISRTIGWMDPFRGWDGSHSAYSTLS 329  
QY 181 DYQCMRRARADKLVLEEMAVPLKSVEDIYERFLYCLKLSKNRPSSELDGQPTKLINE 240  
|:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|  
Db 330 DYQCMRRARADKLVLEEMAVPLKSVEDIYERFLYCLKLSKNRPSSELDGQPTKLINE 389  
QY 241 QLEKINTKGLTINSOPAVNGEKSDSPYVG 270  
|:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|||||||||:|  
Db 390 QLEKINTKGLTINSOPAVNGEKSDSPYVG 419

RESULT 12  
ID AAG36170 standard; Protein; 470 AA.  
XX  
AC AAG36170;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44289.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0123788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.



[illegible]

PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139452
PR	17-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139459
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139463
PR	23-JUN-1999	9905-0140695
PR	24-JUN-1999	9905-0140695
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140991
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141847
PR	01-JUL-1999	9905-0142155
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142920
PR	12-JUL-1999	9905-0143297
PR	13-JUL-1999	9905-0143547
PR	14-JUL-1999	9905-0143624
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144884
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145087
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145089
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145224
PR	23-JUL-1999	9905-0145226
PR	26-JUL-1999	9905-0145276
PR	27-JUL-1999	9905-0145913
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145919
PR	28-JUL-1999	9905-0145951
PR	28-JUL-1999	9905-0146386
PR	02-AUG-1999	9905-0146388
PR	02-AUG-1999	9905-0146389
PR	03-AUG-1999	9905-0147034
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171

```
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.
```

Query Match 54.1%; Score 1184; DB 21; Length 497;  
Best Local Similarity 78.9%; Pred. No. 9.7e-110;

Matches 213; Conservative 28; Mismatches 29; Indels 0; Gaps 0;

```
QY 1 ARVIVTQLFYDIDTFLKFNVDRCRQIGITCPVIGMPTNNYKGFIRMTGCKTKIPADIM 60
DB 214 ADLIVTQLFYDIDTFLKFNVDRCRQIGINCPVIGMPTNNYKGFIRMTGCKTKIPADLT 273
```

```
QY 61 AALEPIKONEEAVKAYGHLGTECKKILAHGKITLHLYTLNMEKSAALAILMNLGILBES 120
DB 274 AALEPIKDNDEAVKAYGHLGTECKKILAHGKITLHLYTLNVDKSYGILMNLGLIDES 333
QY 121 KYSRSLPMWRPAPVRFVEDVPRITFMANRPKSYISRTIGMDQYPRGRKGDSCNPSYCAL 180
DB 334 KISRSLPMWRPAPVRFVEDVPRITFMANRPKSYISRTIGMDQYPRGRKGDSCNPSYCAL 393
QY 181 DYQFMRPRARDKLEVENAVPLKSVEDYERPRRLCYLGLKRSNPMSELDGLOPEKTIINE 240
DB 394 DYQFMRPRARDKLEVENAVPLKSVEDYERPRRLCYLGLKRSNPMSELDGLOPEKTIINE 453
QY 241 QLEKINTKGFLLTINSQAPVNGEKSDSPVVG 270
DB 454 QLEKINTKGFLLTINSQAPVNGEKSDSPVVG 483

RESULT 14
AAR88358 standard; Protein; 656 AA.
AAR88358;
AC AAR88358;
DE 15-OCT-1995 (first entry)
DF
DX
DY
DE Human methylene-tetrahydrofolate-reductase.
KW Methylenetetrahydrofolate-reductase; MTHFR; gene therapy;
KW cardiovascular disease; neurological disease; folic acid metabolism;
KW EC-1.5.1.20; enzyme.
OS Homo sapiens.
PN WO9533054-A1.
PD 07-DEC-1995.
PE 25-MAY-1995; 95WO-CA00314.
PR 26-MAY-1994; 94GB-0010620.
PA (UYMC-) UNIV MCGILL.
PL Goyette P, Rozen R;
DR WPI; 1996-030565/03.
DR N-PSDB; AAT09689.
PT Human methylene:tetrahydrofolate reductase cDNA probe - for
PT detection of sequence abnormalities in methylene:tetrahydrofolate
PT reductase e.g. in cardiovascular, neurological or folic acid
PT metabolism disorders
XX
XX Disclosure; Fig. 6A-6C; 66pp; English.
XX
XX This is the protein sequence encoding human MTHFR, the gene of
XX which has been localised to chromosome 1p36.3. Deficiencies of
XX this protein may lead to cardiovascular and neurological disorders
XX and disorders influences by folic acid metabolism.
XX
XX Sequence 656 AA;
SQ
```

Query Match 44.8%; Score 979; DB 17; Length 656;  
Best Local Similarity 44.9%; Pred. No. 6e-89;

Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;

```
QY 1 ARVIVTQLFYDIDTFLKFNVDRCRQIGITCPVIGMPTNNYKGFIRMTGCKTKIPADIM 60
DB 222 ADLIVTQLFYDIDTFLKFNVDRCRQIGINCPVIGMPTNNYKGFIRMTGCKTKIPADLT 281
QY 61 AALEPIKONEEAVKAYGHLGTECKKILAHG-IKTLHLYTLNMEKSAALAILMNLGILBEE 119
```

Db 282 DVIEPIKNDAAIRNIGIELAVSLCOELLASGLVPGLFHYTLINREMAATEVLKRLGMWTE 341  
QY 120 SKVSRSLPWRPAPVFRVKEVDRIPIFWANRPKSYISRTIGMDQYHGRMGDSCNPSTGAL 179  
Db 342 DP-RRLPFWALSAHPKRREEDVRPIFWASRPKSYIYRTOEWDEFNPGRWGSSSPAFGEL 400  
QY 180 SDYQ--FMRPRARDKTLVEEAVPLKSVEDIYERFRLXCGKLRSN-----PMSRLDG 230  
Db 401 KDYLFLYLSKSPKKEELKMGEEELTSEASVEFVLYLSEGNRNNGHVTCLPWNDEP 459  
QY 231 LOPETKIINOLEKINTKGFILINSOPAVNGEKSDSPYVGWGGPGGYVYOKAYVEPFCR 290  
Db 460 LAAETSLKEELLRVNRGCIITINSQPINNGKPSDPIVGMGSPSGGYVFOKAYLEFFTSR 519  
QY 291 EKLDAVDKCD-RTSLTYMAVNGKSGMSNGQTDVNAVTVGVPAKEIIQPTIVDPVS 349  
Db 520 ETAEALLQVLKKEELRVNTHLVNKGENTNAPLOPNAVTVGIFPGRRIIOPTIVDPVS 579  
QY 350 FNVWKEDEFELMSRGMASLYPEDEASRKLYEEVGGSHFLVSLVNDY-INGDLFAVFPAD 407  
Db 580 FFMWKEAFALMIERMGKLYEESPSRTIIQYIHDNIFLVNLVNDPPLDNCIMQVVED 638  
RESULT 15  
AAEL2607  
ID AAEL2607 standard; Protein; 656 AA.  
XX  
AC AAEL2607;  
XX  
DT 03-JAN-2002 (first entry)  
XX  
DE Human methylenetetrahydrofolate reductase (MTHFR) protein #2.  
XX  
KW Human: methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;  
KW anticense therapy; EC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;  
KW pancreas; lung; brain; cytosolic; colon carcinoma; breast; colorectal;  
KW neuroblastoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN US2001025030-A1.  
XX  
PD 27-SEP-2001.  
XX  
PF 01-DEC-2000; 2000US-0728910.  
XX  
PR 01-MAR-1999; 99US-0258928.  
XX  
PA (ROZE/) ROZEN R.  
XX (SEKH/) SEKHON J.  
XX  
PI Rozen R, Sekhon J;  
XX  
DR WPI; 2001-638509/73.  
XX  
DR N-PSDB; AAD20453.  
XX  
PT New antisense nucleic acids, which are methylenetetrahydrofolate  
PT reductase inhibitors, useful for treating, stabilizing or preventing  
PT cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or  
PT neuroblastoma  
XX  
PS Disclosure; Fig 6; 68pp; English.  
XX  
CC The invention relates to a non allele-specific antisense nucleic acids,  
CC which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)  
CC nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR  
CC catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to  
CC 5-methylenetetrahydrofolate, a co-substrate for methylation of  
CC homocysteine to methionine. The invention provides potential therapy for  
CC individuals with MTHFR deficiency. The non allele-specific antisense  
CC nucleic acids are useful for treating, stabilizing or preventing cancer,  
CC particularly breast carcinoma, colon carcinoma, colorectal carcinoma,  
CC lung cancer, brain cancer, pancreatic cancer, kidney cell cancer.

CC neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene  
CC therapy. The present sequence is human methylenetetrahydrofolate  
CC reductase (MTHFR) protein #2. Human MTHFR gene is mapped to  
CC chromosome 1p36.3.  
XX  
SQ Sequence 656 AA;  
Query Match 44.8%; Score 979; DB 22; Length 656;  
Best Local Similarity 44.9%; Pred. No. 6e-89;  
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;  
QY 1 ARVITQLEFVTDIFLKFVNDCKRQIGTCPIVPSIMINNYKGFIRMTGCKTKIPADIM 60  
Db 222 ADPIITQLFFPADPFFRNVACTDMGITCPIVPGIFIQGHSRLQVLKLSKLEVPQEK 281  
QY 61 AALPEIKNEPAVKAAGIHGTGEMCKKILAHG-IKTLHLTYLNNKESALAILNMLGIEE 119  
Db 282 DVIEPIKNDAAIRNIGIELAVSLCOELLASGLVPGLFHYTLINREMAATEVLKRLGMWTE 341  
QY 120 SKVSRSLPWRPAPVFRVKEVDRIPIFWANRPKSYISRTIGMDQYHGRMGDSCNPSTGAL 179  
Db 342 DP-RRLPFWALSAHPKRREEDVRPIFWASRPKSYIYRTOEWDEFNPGRWGSSSPAFGEL 400  
QY 180 SDYQ--FMRPRARDKTLVEEAVPLKSVEDIYERFRLXCGKLRSN-----PMSRLDG 230  
Db 401 KDYLFLYLSKSPKKEELKMGEEELTSEASVEFVLYLSEGNRNNGHVTCLPWNDEP 459  
QY 231 LOPETKIINOLEKINTKGFILINSOPAVNGEKSDSPYVGWGGPGGYVYOKAYVEPFCR 290  
Db 460 LAAETSLKEELLRVNRGCIITINSQPINNGKPSDPIVGMGSPSGGYVFOKAYLEFFTSR 519  
QY 291 EKLDAVDKCD-RTSLTYMAVNGKSGMSNGQTDVNAVTVGVPAKEIIQPTIVDPVS 349  
Db 520 ETAEALLQVLKKEELRVNTHLVNKGENTNAPLOPNAVTVGIFPGRRIIOPTIVDPVS 579  
QY 350 FNVWKEDEFELMSRGMASLYPEDEASRKLYEEVGGSHFLVSLVNDY-INGDLFAVFPAD 407  
Db 580 FFMWKEAFALMIERMGKLYEESPSRTIIQYIHDNIFLVNLVNDPPLDNCIMQVVED 638

Search completed: February 14, 2003, 23:07:42  
Job time : 73 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2003, 19:26:12 ; Search time 22 Seconds

(without alignments)  
769.197 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187

Sequence: 1 ARVYVQLFDTDFLKFVN.....VSLVDNDYINGDLFAVADF 408

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1834	83.9	594	1 MTHR_ARATH	O08585 arabidopsis
2	979	44.8	656	1 MTHR_HUMAN	P42898 homo sapien
3	943	43.1	654	1 MTHR_MOUSE	Q9WU20 mus musculu
4	895	40.9	603	1 MTHR_SCHPO	Q10258 schizosacch
5	887	40.6	663	1 MTHR_CAEEL	Q17693 caenorhabdi
6	778	35.6	599	1 MTHR_YEAST	P53128 saccharomyc
7	629	28.8	641	1 MTHR_SCHPO	O79927 schizosacch
8	578.5	26.5	657	1 MTHR_YEAST	P46151 saccharomyc
9	222	10.2	296	1 MTHR_YEAST	O67422 aquifex aeo
10	193	8.8	307	1 MTHR_STRLI	O54235 streptomyc
11	189	8.6	292	1 MTHR_HAEIN	P45208 haemophilus
12	189	8.6	296	1 MTHR_ECOLI	P00394 escherichia
13	184	8.4	296	1 MTHR_SALTY	P11003 salmonella
14	176	8.0	298	1 MTHR_ERWCA	P71319 erwina car
15	168	7.7	292	1 MTHR_BUCAL	P57154 buchnera ap
16	105.5	4.8	2156	1 RPI_HUMAN	P56715 homo sapien
17	105	4.8	684	1 YMW5_YEAST	Q04779 saccharomyc
18	101	4.6	548	1 PHRI_CANAL	P43076 candida alb
19	99	4.5	422	1 NCAP_VSVTG	P11212 vesicular s
20	98.5	4.5	1067	1 NCAP_VSVTG	O8XHB3 clostridium
21	97	4.4	422	1 NCAP_VSVST	P03521 vesicular s
22	96	4.4	900	1 SYA_MYCPN	P75368 mycoplasma
23	96	4.4	924	1 CSKP_MOUSE	O70589 mus musculu
24	95	4.3	1005	1 VIA_PSVJ	P28726 peanut stru
25	94	4.3	921	1 CSKP_HUMAN	O1936 homo sapien
26	93.5	4.3	437	1 KDA_CHLN	Q46222 chlamydia p
27	92.5	4.2	1231	1 BLM_CAEEL	O18017 caenorhabdi
28	92	4.2	879	1 PHSG_AOUAE	O69312 aquifex aeo
29	91.5	4.2	879	1 MAMB_CAPII	O95327 capri hircu
30	91	4.2	886	1 DSCI_MOUSE	P55849 mus musculu
31	91	4.2	1026	1 MYIB_DROME	O29979 drosophila
32	91	4.2	1161	1 DP3A_AOUAE	O61125 aquifex aeo
33	90.5	4.1	434	1 PNCB_RHIO	O98d24 rhizobium 1

34	90.5	4.1	568	1 PPO1_AGABI	O00024 agaricus bl
35	90	4.1	391	1 POLG_CYVV	P20177 clover yell
36	90	4.1	549	1 EPD2_CANMA	O74137 candida mal
37	89.5	4.1	387	1 RH18_SCHPO	O74747 schizosacch
38	89.5	4.1	639	1 CAIC_RABIT	Q28902 oryctolagus
39	89.5	4.1	3119	1 CAIC_MOUSE	Q06847 mus musculu
40	89	4.1	446	1 Y612_MERJA	O58029 methanococ
41	89	4.1	494	1 SYK_BACST	O9rtv9 bacillus st
42	89	4.1	610	1 LKHA_MOUSE	P24527 mus musculu
43	88.5	4.0	572	1 SYM_AERPE	O9cy3 aeropyrum p
44	88.5	4.0	1466	1 SPA2_YEAST	P23201 saccharomyc
45	88	4.0	909	1 CSKP_RAT	O62915 rattus norv

## ALIGNMENTS

RESULT 1  
MTHR\_ARATH STANDARD: PRT: 594 AA.  
ID O08585:09SEP59:Q9AJZ1:Q9SU0;  
AC 15-JUN-1999 (Rel. 38, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Methylenetetrahydrofolate reductase (EC 1.5.1.20) (MTHFR2).  
GN MTHFR1 OR ATMG44160 OR F6E13.29.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
ON NCBI\_TaxID=3702;  
OX [1]  
RP SEQUENCE FROM N.A.  
RA Ravanell S., Rebelle F., Douce R.;  
RT "Folate metabolism in higher plants: cloning of a cDNA for  
RT 5,10-methylenetetrahydrofolate reductase in Arabidopsis thaliana.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=20062814; PubMed=10593891;  
RA Roje S., Wang H., McNeil S.D., Raymond R.K., Appling D.R.,  
RA Shachar-Hill Y., Boehrert H.J., Hanson A.D.;  
RT "Isolation, characterization, and functional expression of cDNAs  
RT encoding NADH-dependent methylenetetrahydrofolate reductase from  
RT higher plants.";  
RL J. Biol. Chem. 274:36089-36096(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA STRAIN=cv. Columbia;  
RC MEDLINE=20083487; PubMed=10617197;  
RX Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldguy T.V.,  
RA Buell K.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,  
RA Motil K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayam L.,  
RA Tallon L.J., Gill J.E., Adams M.D., Carreira A.J., Creasy T.H.,  
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,  
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
RA Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
RT thaliana.";  
RL Nature 402:761-768(1999).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA STRAIN=cv. Columbia;  
RC Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;  
RT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the  
RT SSP consortium (Salk/Stanford/PGEC).";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-  
CC methylenetetrahydrofolate + NADPH.  
CC -1- COFACTOR: FAD (By Similarity).  
CC -1- PATHWAY: Folate metabolism.  
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE

```
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; AJ245414; CAB53783.1; -
DR EMBL; AF181967; AAD55788.1; -
DR EMBL; AC004005; AAC23420.2; -
DR EMBL; AY050434; AAK91450.1; -
DR EMBL; AF370515; AAK43892.1; -
DR HSSP; P00394; 1B5T.
DR InterPro; IPR004621; Fadh2_euk.
DR InterPro; IPR003171; Methylrofl_reductase.
DR Pfam; PF02219; MTHR.1.
DR TIGRFAMs; TIGR00677; fadh2_euk; 1.
KM Oxidoreductase; Flavoprotein; FAD; NADP.
FT CONFLICT 72 N -> S (IN REF. 1).
FT CONFLICT 115 P -> H (IN REF. 4; AAK43892).
FT CONFLICT 508 A -> V (IN REF. 1).
SQ SEQUENCE 594 AA; 66802 MW; E2F4F919FCE10EF8 CRC64;

Query Match 83.9%; Score 1834; DB 1; Length 594;
Best local similarity 82.1%; Pred. No. 1.2e-141;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

QY 1 ARVITQLFYPDIDFLKFNDCRQIGTICPIPGIMPINNYKGFIRMFGFCKTIPADIM 60
DB 187 ADLIYTOLEFYDIDFLKFNDCRQIGTICPIPGIMPINNYKGFIRMFGFCKTIPAEVM 246
QY 61 AALEPIKNEAAYKAYGHGHTGEMCKTLAIGITLHLYTNMKSALAIIMNLGIEES 120
DB 247 AALEPIKNEAAYKAYGHGHTGEMCKTLAIGITLHLYTNMKSALAIIMNLGMDIES 306
QY 121 KYSRSLPMPRPANVRYVEDRPIFMANRPKYSRTIGMDQYPHGRMGDCSNPYGALS 180
DB 307 KISRSLPMPRPANVRYVEDRPIFMANRPKYSRTIGMDQYPHGRMGDCSNPYGALS 366
QY 181 DYQMRPAPARCKLVEEMAVPLKSVEDIYERFLYCLGKLSNPMSELGIQPTKTIINE 240
DB 367 DHOEPRPAPARCKLQOEAVVPLKSVEDIQEKFKELGKLSNPMSELGIQPTKTIINE 426
QY 241 QLEKNTGELTINSQPAVNGEKSDPTVWGSGGCVYQKAYVEFFCSKEDALVDKC 300
DB 427 QLIKNGSGELTINSQPAVNGEKSDPTVWGSGGCVYQKAYVEFFCSKEDAVVEKC 486
QY 301 KDRTSLTYMAVNAKDGSMKSNVGTDVNAVTVGVPFAKEIIOPTIVDPVSFVWMDAEFEI 360
DB 487 KALPSITVMAVNAKDGSMKSNVGTDVNAVTVGVPFAKEIIOPTIVDPVSFVWMDAEFEI 546
QY 361 WSRGWASLIYPDEASRKLVVEVGGSHLVLVVDNDYINGDLFAVFAAD 407
DB 547 WSRGWASLIYPDEASRKLVVEVGGSHLVLVVDNDYINGDLFAVFAAD 593

RESULT 2
MTHR_HUMAN STANDARD; PRT; 656 AA.
AC P42898; Q9UOR2;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylentetrahydrofolate reductase (EC 1.5.1.20).
GN MTHR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE OF 1-415 FROM N.A., AND VARIANT GLN-157.
RC TISSUE-Liver;
RX MEDLINE=95004587; PubMed=7920641;
RA Goyette P., Sumner J.S., Milos R., Duncan A.M., Rosenblatt D.S.,
RA Matthews R.G., Rozen R.;
RT "Human methylentetrahydrofolate reductase: isolation of cDNA,
RL mapping and mutation identification.";
RL Nat. Genet. 7:195-200(1994).
RN [12]
RP ERRATUM.
RX MEDLINE=95038846; PubMed=7951330;
RA Goyette P., Sumner J.S., Milos R., Duncan A.M., Rosenblatt D.S.,
RA Matthews R.G., Rozen R.;
RL Nat. Genet. 7:551-551(1994).
RN [13]
RP SEQUENCE FROM N.A.
RX Rozen R., Goyette P.;
RT "cDNA for human methylentetrahydrofolate reductase.";
RL Patent number WO9533054, 07-DEC-1995.
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=98345426; PubMed=9680386;
RA Goyette P., Pal A., Milos R., Frosst P., Tiran P., Chen Z., Chan M.,
RA Rozen R.;
RT "Gene structure of human and mouse methylentetrahydrofolate reductase
RL (MTHR).";
RL Mamm. Genome 9:652-656(1998).
RN [15]
RP SEQUENCE FROM N.A.
RX Homberger A., Lindebank M., Winter C., Rapp B., Koch H.G.;
RT "Revised translation initiation site of the human
RL methylentetrahydrofolate reductase (MTHR).";
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [16]
RP VARIANTS Q-52; M-227; L-251; C-325; C-335 AND C-357, AND REVISION TO
RL 177.
RX MEDLINE=95243219; PubMed=7726158;
RA Goyette P., Frosst P., Rosenblatt D.S., Rozen R.;
RT "Seven novel mutations in the methylentetrahydrofolate reductase
RL gene and genotype/phenotype correlations in severe
RL methylentetrahydrofolate reductase deficiency.";
RL Am. J. Hum. Genet. 56:1052-1059(1995).
RN [17]
RP VARIANT VAL-222.
RX MEDLINE=95375773; PubMed=7647779;
RA Frosst P., Blom H.J., Milos R., Goyette P., Sheppard C.A.,
RA Matthews R.G., Boers G.J.H., den Heijer M., Kluitjans L.A.J.,
RA van den Heuvel L.P., Rozen R.;
RT "A candidate genetic risk factor for vascular disease: a common
RL mutation in methylentetrahydrofolate reductase.";
RL Nat. Genet. 10:111-113(1995).
RN [18]
RP VARIANTS PRO-51; PRO-323 AND CYS-377.
RX MEDLINE=97094184; PubMed=8940272;
RA Goyette P., Christensen B., Rosenblatt D.S., Rozen R.;
RT "Severe and mild mutations in cis for the methylentetrahydrofolate
RL reductase (MTHR) gene, and description of five novel mutations in
RL MTHR.";
RL Am. J. Hum. Genet. 59:1268-1275(1996).
RN [19]
RP VARIANT VAL-222.
RX MEDLINE=98213566; PubMed=9545406;
RA Schneider J.A., Rees D.C., Liu Y.-T., Clegg J.B.;
RT "Worldwide distribution of a common methylentetrahydrofolate
RL reductase mutation.";
RL Am. J. Hum. Genet. 62:1258-1260(1998).
RN [110]
RP VARIANT ALA-428.
RX MEDLINE=98213566; PubMed=9545395;
RA van der Put N.M.J., Gabreels F., Stevens E.M.B., Smeltink J.A.M.,
RA Trijbels F.J.M., Eskes T.K.A.B., van den Heuvel L.P., Blom H.J.;
RT "A second mutation in the methylentetrahydrofolate reductase gene: an
RL additional risk factor for neural-tube defects?";
```

RL Am. J. Hum. Genet. 62:1044-1051(1999).  
 RP [11]  
 RP VARIANTS SER-324 AND GLY-339.  
 RX MEDLINE=98454315; PubMed=9781030;  
 RA Kluijtmans L.A.J., Wendel U., Stevens E.M.B., van den Heuvel L.P.W.J.,  
 RT Trijbels F.J.M., Blom H.J.;  
 RA "Identification of four novel mutations in severe  
 RT methylentetrahydrofolate reductase deficiency.";  
 RL Eur. J. Hum. Genet. 6:257-265(1998).  
 RN [12]  
 RP VARIANT ALA-428.  
 RX MEDLINE=98387934; PubMed=9719624;  
 RA Weisberg I., Tran P., Christensen B., Sibani S., Rozen R.;  
 RT "A second genetic polymorphism in methylentetrahydrofolate reductase  
 (MTHFR) associated with decreased enzyme activity.";  
 RL Mol. Genet. Metab. 64:169-172(1998).  
 RN [13]  
 RP VARIANTS ASP-387; LEU-572 AND IYS-586.  
 RX MEDLINE=20145670; PubMed=10679944;  
 RA Sibani S., Christensen B., O'Ferrall E., Saadi I., Hwu-Tim F.,  
 RA Rosenblatt D.S., Rozen R.;  
 RT "Characterization of six novel mutations in the  
 RT methylentetrahydrofolate reductase (MTHFR) gene in patients with  
 RT homocystinuria.";  
 RL Hum. Mutat. 15:280-287(2000).  
 CC -1- FUNCTION: CATALYZES THE CONVERSION OF 5,10-  
 CC METHYLENTERAHDROFOLATE TO 5-METHYLETETRAHYDROFOLATE, A CO-  
 CC SUBSTRATE FOR HOMOCYSTEINE REMETHYLATION TO METHIONINE.  
 CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-  
 CC methylentetrahydrofolate + NADPH.  
 CC -1- COFACTOR: FAD.  
 CC -1- ENZYME REGULATION: ALLOSTERICALLY REGULATED BY S-  
 CC ADENOSYLMETHIONINE.  
 CC -1- PATHWAY: Folate metabolism.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- POLYMERISM: VARIANT VAL-222 IS A COMMON POLYMORPHISM WHICH HAS  
 CC BEEN IMPLICATED IN THREE MULTIFACTORIAL DISEASES: OCCLUSIVE  
 CC VASCULAR DISEASE, NEURAL TUBE DEFECTS (NTD) AND COLON CANCER.  
 CC -1- DISEASE: DEFECTS IN MTHFR ARE THE CAUSE OF HOMOCYSTEINEMIA  
 CC (HOMOCYSTEINURIA II), AN AUTOSOMAL RECESSIVE DISEASE WHICH SHOWS A  
 CC WIDE RANGE OF CLINICAL SYMPTOMS, SUCH AS DEVELOPMENTAL DELAY,  
 CC SEVERE MENTAL RETARDATION, PERINATAL DEATH, PSYCHIATRIC  
 CC DISTURBANCES, AND LATER-ONSET NEURODEGENERATIVE DISORDERS.  
 CC -1- SIMILARITY: BELONGS TO THE METHYLENTERAHDROFOLATE REDUCTASE  
 CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U09806; AAA74440.2; -;  
 DR EMBL: A47328; CAA03053.1; -;  
 DR EMBL: AF105987; AAD17965.1; JOINED.  
 DR EMBL: AF105977; AAD17965.1; JOINED.  
 DR EMBL: AF105978; AAD17965.1; JOINED.  
 DR EMBL: AF105979; AAD17965.1; JOINED.  
 DR EMBL: AF105980; AAD17965.1; JOINED.  
 DR EMBL: AF105981; AAD17965.1; JOINED.  
 DR EMBL: AF105982; AAD17965.1; JOINED.  
 DR EMBL: AF105983; AAD17965.1; JOINED.  
 DR EMBL: AF105984; AAD17965.1; JOINED.  
 DR EMBL: AF105985; AAD17965.1; JOINED.  
 DR EMBL: AF105986; AAD17965.1; JOINED.  
 DR EMBL: AF105987; AAD17965.1; JOINED.  
 DR EMBL: AF105988; AAD17965.1; JOINED.  
 DR EMBL: AF105989; AAD17965.1; JOINED.  
 DR EMBL: AF105990; AAD17965.1; JOINED.  
 DR EMBL: AF105991; AAD17965.1; JOINED.  
 DR EMBL: AF105992; AAD17965.1; JOINED.  
 DR EMBL: AF105993; AAD17965.1; JOINED.  
 DR EMBL: AF105994; AAD17965.1; JOINED.  
 DR EMBL: AF105995; AAD17965.1; JOINED.  
 DR EMBL: AF105996; AAD17965.1; JOINED.  
 DR EMBL: AF105997; AAD17965.1; JOINED.  
 DR EMBL: AF105998; AAD17965.1; JOINED.  
 DR EMBL: AF105999; AAD17965.1; JOINED.  
 DR EMBL: AF106000; AAD17965.1; JOINED.  
 DR EMBL: AF106001; AAD17965.1; JOINED.  
 DR EMBL: AF106002; AAD17965.1; JOINED.  
 DR EMBL: AF106003; AAD17965.1; JOINED.  
 DR EMBL: AF106004; AAD17965.1; JOINED.  
 DR EMBL: AF106005; AAD17965.1; JOINED.  
 DR EMBL: AF106006; AAD17965.1; JOINED.  
 DR EMBL: AF106007; AAD17965.1; JOINED.  
 DR EMBL: AF106008; AAD17965.1; JOINED.  
 DR EMBL: AF106009; AAD17965.1; JOINED.  
 DR EMBL: AF106010; AAD17965.1; JOINED.  
 DR EMBL: AF106011; AAD17965.1; JOINED.  
 DR EMBL: AF106012; AAD17965.1; JOINED.  
 DR EMBL: AF106013; AAD17965.1; JOINED.  
 DR EMBL: AF106014; AAD17965.1; JOINED.  
 DR EMBL: AF106015; AAD17965.1; JOINED.  
 DR EMBL: AF106016; AAD17965.1; JOINED.  
 DR EMBL: AF106017; AAD17965.1; JOINED.  
 DR EMBL: AF106018; AAD17965.1; JOINED.  
 DR EMBL: AF106019; AAD17965.1; JOINED.  
 DR EMBL: AF106020; AAD17965.1; JOINED.  
 DR EMBL: AF106021; AAD17965.1; JOINED.  
 DR EMBL: AF106022; AAD17965.1; JOINED.  
 DR EMBL: AF106023; AAD17965.1; JOINED.  
 DR EMBL: AF106024; AAD17965.1; JOINED.  
 DR EMBL: AF106025; AAD17965.1; JOINED.  
 DR EMBL: AF106026; AAD17965.1; JOINED.  
 DR EMBL: AF106027; AAD17965.1; JOINED.  
 DR EMBL: AF106028; AAD17965.1; JOINED.  
 DR EMBL: AF106029; AAD17965.1; JOINED.  
 DR EMBL: AF106030; AAD17965.1; JOINED.  
 DR EMBL: AF106031; AAD17965.1; JOINED.  
 DR EMBL: AF106032; AAD17965.1; JOINED.  
 DR EMBL: AF106033; AAD17965.1; JOINED.  
 DR EMBL: AF106034; AAD17965.1; JOINED.  
 DR EMBL: AF106035; AAD17965.1; JOINED.  
 DR EMBL: AF106036; AAD17965.1; JOINED.  
 DR EMBL: AF106037; AAD17965.1; JOINED.  
 DR EMBL: AF106038; AAD17965.1; JOINED.  
 DR EMBL: AF106039; AAD17965.1; JOINED.  
 DR EMBL: AF106040; AAD17965.1; JOINED.  
 DR EMBL: AF106041; AAD17965.1; JOINED.  
 DR EMBL: AF106042; AAD17965.1; JOINED.  
 DR EMBL: AF106043; AAD17965.1; JOINED.  
 DR EMBL: AF106044; AAD17965.1; JOINED.  
 DR EMBL: AF106045; AAD17965.1; JOINED.  
 DR EMBL: AF106046; AAD17965.1; JOINED.  
 DR EMBL: AF106047; AAD17965.1; JOINED.  
 DR EMBL: AF106048; AAD17965.1; JOINED.  
 DR EMBL: AF106049; AAD17965.1; JOINED.  
 DR EMBL: AF106050; AAD17965.1; JOINED.  
 DR EMBL: AF106051; AAD17965.1; JOINED.  
 DR EMBL: AF106052; AAD17965.1; JOINED.  
 DR EMBL: AF106053; AAD17965.1; JOINED.  
 DR EMBL: AF106054; AAD17965.1; JOINED.  
 DR EMBL: AF106055; AAD17965.1; JOINED.  
 DR EMBL: AF106056; AAD17965.1; JOINED.  
 DR EMBL: AF106057; AAD17965.1; JOINED.  
 DR EMBL: AF106058; AAD17965.1; JOINED.  
 DR EMBL: AF106059; AAD17965.1; JOINED.  
 DR EMBL: AF106060; AAD17965.1; JOINED.  
 DR EMBL: AF106061; AAD17965.1; JOINED.  
 DR EMBL: AF106062; AAD17965.1; JOINED.  
 DR EMBL: AF106063; AAD17965.1; JOINED.  
 DR EMBL: AF106064; AAD17965.1; JOINED.  
 DR EMBL: AF106065; AAD17965.1; JOINED.  
 DR EMBL: AF106066; AAD17965.1; JOINED.  
 DR EMBL: AF106067; AAD17965.1; JOINED.  
 DR EMBL: AF106068; AAD17965.1; JOINED.  
 DR EMBL: AF106069; AAD17965.1; JOINED.<

Query Match	Best Local Similarity	Matches 188;	Conservative 79;	Mismatches 138;	Indels 14;	Gaps 7;
QY 1	ARVYITQLFYDDIFLKEVYNDCRQIGITCPIYPGIMPINNYGFLRMGFCCTKIPADIM 60	1	1	1	1	1
DB 222	ADFIITQLEFEDADTEFRFVKACTDMDGICITCPIYGFPIIGIGYSLQLVRLSKLVEPOEIK 281	1	1	1	1	1
QY 61	ADLEPKDNEEAVKRYAGHGLGTEMCKKILAHG-ITKIHLYTLNMRKSAALAIINGLIDE 119	1	1	1	1	1
DB 282	DVIEPKDDAKDAIRNGIELANVSLCOELASGLVGLHLYTLINREMAITTEVLRUGMTE 341	1	1	1	1	1
QY 120	SKVRSLSLPMRRRANRYEVDVPIFWMARPKSYISRTIGMDQYPHGRGDCSNPSYGL 179	1	1	1	1	1
DB 342	DF-RRPLPALNKKRRREDVRIPIFMSRSPKSYIRQEMWELFPGNRGNSSPAFGL 400	1	1	1	1	1
QY 180	SDYQ-EMRRARADKKLYEEMAVPLKSVEDIYERFRLYCGLKRSN-----PMSELDG 230	1	1	1	1	1
DB 401	KDYLYLFYTKSKSPKBEILKMGEEELTSEASVEVEFVLYLSGPNRNGHKVTCPLMND-EP 459	1	1	1	1	1
QY 231	LOPEKRIINEOLEKINTKGFLLINQPAVNGEKSDSPYGMGPGGYVQAKAYVEFFCK 290	1	1	1	1	1
DB 460	LAETSLSLKEELLRRNRQGIILINQPNINGRPSSDPIYGMWPGSGGYVQAKAYVEFFTSR 519	1	1	1	1	1
QY 291	EKLDAIYDCKD-RSLTGMAYVKNQSGMSNGQDVNAVGVNGVPPAKETIOPITYDPS 349	1	1	1	1	1
DB 520	EFAEALLQYLKLYELRVNHYLVNKGENTINAPLQPNAYVNGIRFGRITIDPTVYDPS 579	1	1	1	1	1
QY 350	FVWVDEAFEISRGWASLYPDEASRKLVEYVGSGHFLVSLVNDY-INGDLFAVFAD 407	1	1	1	1	1

Db 580 FFWKDEAFALMIERMGKLYEESPSRTIIQYIHNDNYFLVNLVNDPELDNCIMQVED 638

RESULT 3

MTNR\_MOUSE STANDARD; PRT; 654 AA.

ID MTNR\_MOUSE

AC 09JUN20;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Methylene-tetrahydrofolate reductase (EC 1.5.1.20).

GN MTNR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98345426; PubMed=9680386;

RA Goyette P., Pal A., Milos R., Frost P., Tran P., Chen Z., Chan M.,

RA Rozen R.;

RT "gene structure of human and mouse methylenetetrahydrofolate reductase (MTHFR).";

RL Mamm. Genome 9:652-656(1998).

CC CC

CC FUNCTION: CATALYZES THE CONVERSION OF 5,10-METHYLENETHETRAHYDROFOLATE TO 5-METHYLETETRAHYDROFOLATE, A CO-SUBSTRATE FOR HOMOCYSTEINE REMETHYLATION TO METHIONINE.

CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-methylenetetrahydrofolate + NADPH.

CC -1- COFACTOR: FAD (BY SIMILARITY).

CC -1- ENZYME REGULATION: ALLOSTERICALLY REGULATED BY S-ADENOSYLMETHIONINE (BY SIMILARITY).

CC -1- PATHWAY: Folate metabolism.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE METHYLENETHETRAHYDROFOLATE REDUCTASE (EC 1.5.1.20/EC 1.7.99.5) FAMILY.

CC

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL: AF105988; AAD20313.1; -.

DR EMBL: AF105989; AAD20313.1; JOINED.

DR EMBL: AF105989; AAD20313.1; JOINED.

DR EMBL: AF105990; AAD20313.1; JOINED.

DR EMBL: AF105991; AAD20313.1; JOINED.

DR EMBL: AF105992; AAD20313.1; JOINED.

DR EMBL: AF105993; AAD20313.1; JOINED.

DR EMBL: AF105994; AAD20313.1; JOINED.

DR EMBL: AF105995; AAD20313.1; JOINED.

DR EMBL: AF105996; AAD20313.1; JOINED.

DR EMBL: AF105997; AAD20313.1; JOINED.

DR HSSP: P00394; 1B5T.

DR MGD: MGI:106639; Mthfr.

DR InterPro: IPR004621; Fadh2\_euk.

DR InterPro: IPR003171; Mchdrof\_redctse.

DR Pfam: PF02219; MTNR\_1.

DR TIGRFAMs: TIGR00677; fadh2\_euk; 1.

KW OXidoreductase; FlavoProtein; FAD; NADP.

SQ SEQUENCE 654 AA; 74649 MW; 12AD31806B371E17 CRC64;

Query Match 43.1%; Score 943; DB 1; Length 654;

Best Local Similarity 43.0%; Pred. No. 4,5e-69;

Matches 180; Conservative 84; Mismatches 141; Indels 14; Gaps 7;

QY 1 ARVLTQLEFDTDIFLKFVNDCKRGITGTCIVGIMINNYKFFIMTGCKTKIPADIM 60

Db 221 ADFIITQLFPEASTFFSFVAKCEIGISCIPLGIFPIQGTJSLRQLVMSKLEVFQKIK 280

QY 61 AALEPIKDNDEAVKAYGHIHGTEMCKTILAHG-ITKTLHYLTLMKESALATIMNLGILEE 119

Db 281 DVEIEPKNDAAIRNNGTIGELAVRLCRELDLSGLVGHEFTITNRVATMEYVKQLGMMTE 340

QY 120 SKVRSRLPWRBRPAPVFRKEDVRPIFWANRPKSYISRTIGMDQYDHGKMGSCNPSYAL 179

Db 341 DP-RRLPALSAHPKRRREDVRIPIFWASRPKSYIYRIQDMDEFPNGRMGNSSPAFCGL 399

QY 180 SDYQ--FMPPRRADKKLYEEMAVPLKSVEDIYERRLYCLG-----KLSNPSELDG 230

Db 400 KDYYLFYLYKSSPRELLKMGEEELTSESVYEVEHNLGSPNNGRYVCLPLMND-EP 458

QY 231 LQPEFKINEOLEKINTKGFITINSOPAVNGKSSPFGYMGCGPGGYOKAYVEFFCSK 290

Db 459 LAAETSLAMEELLRRNRGILITINSQPINAKPSSDPVYMGSPSGYFQKAYLEFFTSR 518

QY 291 EKLDALVDKCD-RFSLYMAVANKDSMKSNVQGTDVNAVTVGWYPAKEIIQPTIVDPS 349

Db 519 ETVEALLQVLTLYELRVNHYIDVAGENTNAPELQPNNAVTVGIPRGREIIQPTVVDPS 578

QY 350 FFWKDEAFETSRGMASTYRPEDEASKLYEEVGGSHFLVSLVNDY-INGDLFVAFVD 407

Db 579 FFWKDEAFALMIERMGKLYEESPSRTIIQYIHNDNYFLVNLVNDPELDNCIMQVED 637

RESULT 4

MTNR\_SCHPO STANDARD; PRT; 603 AA.

ID MTNR\_SCHPO

AC Q10258;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Probable methylenetetrahydrofolate reductase 1 (EC 1.5.1.20).

GN SPAC56P8.10.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgourou K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tvey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Voicakert G., Aert R., Robben J., Grymoprez B.,

RA Weltjens I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hubert H.,

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Punnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaude V., Mottler S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Galliard C., Tallade V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerruti L., Lowe T., McComb W.R., Paulsen I., Potashkin J.,

RA Stpokovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe.";

RL Nature 415:871-880(2002).

CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-methylenetetrahydrofolate + NADPH.

CC -1- COFACTOR: FAD (BY SIMILARITY).

CC -1- PATHWAY: Folate metabolism.

-1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE (EC 1.5.1.20/EC 1.7.99.5) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z69728; CAA93561.1; -  
DR HSP: P00394; 1B5T.  
DR InterPro: IPR004621; Fadh2\_euk.  
DR Pfam: PF02219; MTHFR; 1.  
DR TIGRFAMs: TIGR00677; fadh2\_euk; 1.  
KM Oxidoreductase; Flavoprotein; FAD; NADP.  
SQ SEQUENCE 603 AA; 69012 MW; 38519FEF7830331P CRC64;

Query Match 40.9%; Score 895; DB 1; Length 603;  
Best Local Similarity 45.2%; Pred. No. 3.3e-65;  
Matches 189; Conservative 57; Mismatches 148; Indels 24; Gaps 8;

OY 1 ARVITQLEFDYDIDFLKFNDCROIGTICPIYVGPIMPINNTKGFTRMTGFECKTKIPADIM 60  
DB 178 ADFIVTQMEFYDVDFIYAWDKVRAAGINIPFPGIMPIDAMDSFTRRAKWSGVKIPQHFH 237

OY 61 AALEPIKNEEAVKAYGIIHGTGEMCKKILAHG-IKTLHLTYLTMKESALAILMNLGLTES 120  
DB 238 DFLVVKDDDEGVRRGVELIYEMKRLIASITRLHLYTMLEKAVAKMIIERGLDEN 297

OY 121 KYSRSLPMRRPANNV-----FVKKEDVRFIFWANNPKSYISRTIGMDYPHGRWDS 171  
DB 298 IA----PYDTNNVELJTNASSODRINRGVRIEWRTRNESYVSTQWDELPHGRWDS 353

OY 172 CNPSGALSDYQF---MRPRADKLVEMAVPLKSVSDIYERRLYGLKLRNPSSEL 228  
DB 354 RSPAGEEDAIRGLMSP---KEITTSWGSF-KSYSEIDLPARCEKRISSLPMSDL 408

OY 229 DGLQPETKIINEOLEKINTKGFLLTNSQPAVNGEKSDSPYVGWGPGGYVOKAYVEFC 288  
DB 409 P-ISEADLIRQQLSMNRNFLTINSOPALNGEKSSHPVGWGPFPNGYVOKAYVEFV 467

OY 269 SKEKIDALVDCCKDRTSLTYMAVNGKSGWSKNSGQTDVNAVTVGVPKELIOTPIYDPV 348  
DB 468 HPSLINELKEFYVKKLNSYSYFVTNKGDLDTNSQEIYPNAVTVGVPKELIOTPIYEST 527

OY 349 SFNNWKDAFEIWSRGASLYPDEASRKLYEEVGSGLFVSLVNDYING-DLFAVNF 405  
DB 528 SFLAWKDEAYSL-GMEWANNAYSPDISISKRLLVSMKKEWFLCVIYDNDFOGQSLDFVF 584

RESULT 5  
MTHR\_CAEEL  
ID MTHR\_CAEEL STANDARD; PRT; 663 AA.  
AC Q17693;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable methylenetetrahydrofolate reductase (EC 1.5.1.20).  
GN C06A8.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderae; Caenorhabditis.  
ON NCBI\_TaxID=6239;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Bristol N2;  
RA Leimbach D.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP REVISIONS.

RA Waterston R.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADPH.  
CC methylenetetrahydrofolate + NADPH.  
CC -1- COFACTOR: FAD (By Similarity).  
CC -1- PATHWAY: Folate metabolism.  
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE  
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U39849; AAA81048.2; -  
DR HSP: P00394; 1B5T.  
DR WormPep: C06A8.1; CE30593.  
DR InterPro: IPR004621; Fadh2\_euk.  
DR InterPro: IPR003171; Methylot\_redctse.  
DR Pfam: PF02219; MTHFR; 1.  
DR TIGRFAMs: TIGR00677; fadh2\_euk; 1.  
KM Oxidoreductase; Flavoprotein; FAD; NADP.  
SQ SEQUENCE 663 AA; 75486 MW; EA94806A2C3BCICD CRC64;

Query Match 40.6%; Score 887; DB 1; Length 663;  
Best Local Similarity 44.0%; Pred. No. 1.7e-64;  
Matches 182; Conservative 75; Mismatches 135; Indels 22; Gaps 10;

OY 1 ARVITQLEFDYDIDFLKFNDCROIGTICPIYVGPIMPINNTKGFTRMTGFECKTKIPADIM 60  
DB 235 ANFYITQLEFAEFEFKFDYDCREIGTITPITGIMPIMKYESIKRIAKISQLEIPIHIL 294

OY 61 AALEPIKNEEAVKAYGIIHGTGEMCKKILAHG-IKTLHLTYLTMKESALAILMNLGL--I 117  
DB 295 DDLPIKHDDDAVQKYGTERCIEKCRLLDNGTAPSIHLIYTNMREGSIREILKSLGLMLT 354

OY 118 EESKVSRLSPWR-RPANNVRFKEDVRFIFWANNPKSYISRTIGMDYPHGRWDSNPSY 176  
DB 355 EGDV--FPWKNRSOPHRCLESVRPIYSPFRSRYITRTWMDQFPGNRWGNSSSPA 411

OY 177 GALS DY---QFMRPRADKLVEMAVPLKSVSDIYERFRLYGL-----KLRSNPS 226  
DB 412 GDVSSYLSLTLTVRNADRLA-MFGANITSEFDVAVRVPLINTQAPNADGVAKVYLPMT 470

OY 227 ELD-GLQPETKIINEOLEKINTKGFLLTNSQPAVNGEKSDSPYVGWGPGGYVOKAYVE 285  
DB 471 EAEFGVQPEFSLISEQLWMCNENGILTNSQSPVNGAPSTDPVVGKPGGYCYOKAYLE 530

OY 286 FPCSKEKIDALVDCCKDR--TSLTYMAVNGKSGW-KSNVGQTDVNAVTVGVPKELIOT 342  
DB 531 CFMTAELSDKLIQITIEEPVVRVYHAINKDSFTEDTTPALNAVTVGVPKELIOT 590

OY 343 TIYDVPVSFNNWKDAFEIWSRGASLYPDEASRKLYEEVGSGLFVSLVNDY 396  
DB 591 TVYDPLPSFRAMRDEAYQMMAGQDGYPRKESKSYGIAKVHDEFRLVTLVND 644

RESULT 6  
MTHS\_YEAST  
ID MTHS\_YEAST STANDARD; PRT; 599 AA.  
AC P53128; Q92318;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Methylenetetrahydrofolate reductase 2 (EC 1.5.1.20).  
GN MET13 OR MET11 OR YGL125W OR G2882.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
ON NCBI\_TaxID=4932;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051592; PubMed=8896269;  
 RA Ticon B., Rodriguez-Torres A.M., Rodriguez-Belmonte E., Cadahia J.L.,  
 RA Cerdan E.,  
 RT "Identification of a putative methylenetetrahydrofolate reductase by  
 RT sequence analysis of a 6.8 kb DNA fragment of yeast chromosome VII.",  
 RL Yeast 12:1047-1051(1996).  
 CC -1 CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-  
 CC methylenetetrahydrofolate + NADPH.  
 CC -1 COFACTOR: FAD (By SIMILARITY).  
 CC -1 PATHWAY: Folate metabolism.  
 CC -1 SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE  
 CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: 272647; CA63833.1; -  
 CC DR EMU: X94106; CA63833.1; -  
 CC DR HSSP: P00394; 1B5T.  
 CC DR SGD: S0003093; MET13.  
 CC DR InterPro: IPR004621; Fadh2\_euk.  
 CC DR InterPro: IPR003171; Methydrf\_redctse.  
 CC DR Pfam: PF02219; MTHFR.1.  
 CC DR TIGRFAMS: TIGR00677; fadh2\_euk.1.  
 CC Oxidoreductase: Flavoprotein; FAD; NADP.  
 CC CONFLICT 73 73 A -> R (IN REF. 1; CA63833).  
 CC FT SEQUENCE 599 AA; 68472 MW; AD3465B52A4E700 CMC64;  
 SQ

Query Match 35.6%; Score 778; DB 1; Length 599;  
 Best local similarity 40.9%; Pred. No. 1,1e-55;  
 Matches 168; Conservative 76; Mismatches 147; Indels 20; Gaps 11;

QY 4 ITYQLFYDIDILKLVNDCROGICPIYVGIIMPINNNKGFIRMGFCFKTRPADIMAL 63  
 DB 185 ITYQFYDIDILKLVNDCROGICPIYVGIIMPINNNKGFIRMGFCFKTRPADIMAL 63  
 QY 64 EPIKDNEAVKAYGHLGEMCKKILAHG-ITLHLTYLNMESALAILMNLGLI-ESRK 121  
 DB 245 DPKDDELVRIGTLIYEMCKKILAHG-ITLHLTYLNMESALAILMNLGLI-ESRK 121  
 QY 122 VSRH---LPWRRPAAVFEKEDVREIFMANRPKSTISRTIGW--DQYPHGMGDSGNPS 175  
 DB 305 FNAHPLAVLPWRKSLNPKRKNNEVRIEFWRKRRYSYVARTSQMAVDEPNGRFGDSSSPA 364  
 QY 176 YGLA-----SDYQPMRRADKTLVEMAVPLKSVEDIYRFRFLYCLGKRSNPSSELQGL 231  
 DB 365 FGDLGCGSD--LIRSA--NRCLELWSP--TSINDVAFVLYNLGNLKLCPMSDIP-I 418  
 QY 232 QPRTKIINOLEKINTKGFLLTINSOPAVNGEKSDSPFVGWGGPGVYVYOKAVVEFGCSKE 291  
 DB 419 NDEINIKAKHLIELNOSHITITINSOPAVNGEKSDSPFVGWGGPGVYVYOKAVVEFGCSKE 291  
 QY 292 KIDALVCKDKRTSLTYMAVNNKDSKSN-VGQTDVNAVTVGVPKAKELIQPTIYDVPSF 350  
 DB 479 KLPKRLIDTLKNNFLLFYAIDSGDLSNHPNDSKSNNAVTVGVPKAKELIQPTIYDVPSF 350  
 QY 351 NWKDAEFLWISRGASLIPDEDA-SRKIVVEVGSHPFLVSYVDNDYINGD 400  
 DB 539 LAKKEFYHLINEMKLMNKKYDKPHSAQPLQSLIDDYCLAVINDYISPD 589

RESULT 7  
 MTHS\_SCHPO STANDARD: PRT: 641 AA.  
 AC 074927; P78770; Q9UT80;  
 DT 15-JUL-1999 (Rel. 38, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Methylenetetrahydrofolate reductase 2 (EC 1.5.1.20).  
 GN MTHFR2 OR SPAC343.10.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=968;  
 RA Naula N.;  
 RL Thesis (1998), University of Bern, Switzerland.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Godle A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,  
 RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkstree G., Aert R., Robben J., Grymoprez B.,  
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabell C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehner H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purrelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motter S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., Mccombie W.R., Paulsen I., Potashkin J.,  
 RA Shipakovsky G.V., Ussery D., Barrett B.G., Nurse P.;  
 RA "The genome sequence of Schizosaccharomyces pombe".  
 RL Nature 415:871-880(2002).  
 RN [3]  
 RP SEQUENCE OF 11-641 FROM N.A.  
 RC STRAIN=PR745;  
 RX MEDLINE=9816722; PubMed=9501991;  
 RA Yoshiooka S., Kato K., Nakai K., Okayama H., Nojima H.;  
 RT "Identification of open reading frames in Schizosaccharomyces pombe  
 RT cDNAs.";  
 RL DNA Res. 4:363-369(1997).  
 CC -1 CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-  
 CC methylenetetrahydrofolate + NADPH.  
 CC -1 COFACTOR: FAD (By SIMILARITY).  
 CC -1 PATHWAY: Folate metabolism.  
 CC -1 SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE  
 CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: AJ011686; CA09738.1; -  
 CC DR EMBL: AL109739; CAB52273.1; -  
 CC DR EMBL: D89118; BAA13780.1; -  
 CC DR HSSP: P00394; 1B5T.  
 CC DR InterPro: IPR004621; Fadh2\_euk.  
 CC DR InterPro: IPR003171; Methydrf\_redctse.

```

DR Pfam; PF02219; MTHFR; 1.
DR TIGRFAMs; TIGR00677; fadh2_euk; 1.
KW Oxidoreductase; Flavoprotein; FAD; NADP.
FT CONFLICT 202 202 E -> D (IN REF. 3).
FT CONFLICT 281 281 A -> A (IN REF. 3).
FT CONFLICT 455 455 A -> V (IN REF. 3).
FT CONFLICT 537 641 QVTVYAGNNKSEFLTNPKKGASAVTWGYPGREITOSTII
AEVSEKAMLSSEFQWGMENMAYNSKNSPKRLLENCINDRM
LVTVIHDFMDKNGKLVLELDF -> HVSNSFSSISDST
CFORVSKPLERIRARADRSKFSVQK (IN REF. 3).
FT CONFLICT
FT SEQUENCE 614 614 I -> L (IN REF. 1).
SO SEQUENCE 641 AA; 72140 MW; E7AFDDJF2CIC8A4 CRC64;

Query Match
Best Local Similarity 32.2%; Pred. No. 1.6e-43;
Matches 147; Conservative 80; Mismatches 161; Indels 68; Gaps 13;

QY 1 ARVIVTQLFVDTDFLFKFNVDRCRQIGTC---PIVGPIMINNYKGFIRMTCCKTKIP 56
D 184 ADFITQIFVEPEAFIKFENFVRNHSNALRNIPPIPAIMPISQVSLKMTKRLCGCSVP 243
QY 57 ADIMALEPIDNEAVAKYGIHGTCEMKKIL-----ANGITLHLYTNMKSALAI 110
D 244 SLMQRLNAAPDDDAIKNIGVEHLVDMIKKIMNVQGRVHG---FHECTLNLERSVALI 300
QY 111 LMLGLI-----EESKVSRS-----LPMRRP-----ANVFR 136
D 301 LKNSGLLTKRKQVSEMEDEKILRTTKRLSLDEPAHLNQQVYVPSQAPADSSNLIFV 360
QY 137 VKE-----DVRPIFWANPKYSISRTIGMDQYPRGRWGDSGNPSYALSDY-- 182
D 361 TSKQSVSGHKNDLNEEAP-FVSESGSGVLGRQAMNDQFTGRGRDPSPAYGEIDGYP 419
QY 183 QEMPRRADKTLVEEMNAPPLKSVEDITYRRFLYCLGKLRSPNSELDELQEPETKINEL 242
D 420 TLHPPPSALKL--WGYPVE-SDITSLEFKHIMSDISALPMD-EEVEVETITKYL 474
QY 243 EKINTKGLTINSOPAVANGESKSDSPYVGMGGPGGYVYOKAVEFFCSKEKIDALVDCKD 302
D 475 LKLNQNSMWTYGSQPAVANGASADVPFGMGKGRVFPKAVEGCFVANKDKLDITKMD 534
QY 303 RLSLTLYAVNKDGSKSNVSGTQDVNAVTVGPAKEIIIOPTIVDPVSEFNWKDEAFETWS 362
D 535 NQVTVYVYAGNNKSEFLTNAPKDGASAVTWGYPGREITOSTIIAEVSEKAMLSSEFQW 594
QY 363 RGMASLYPEDEASRKLVEEVSHTFLVLYNDYIN 398
D 595 E-WANLYSKNTPSKRLLENCINDRMVLTVIHDFMD 629

RESULT 8
MTHR_YEAST STANDARD: PRT: 657 AA.
AC P46151;
DT 01-NOV-1995 (Ref. 32, Created)
DT 01-OCT-1996 (Ref. 34, Last sequence update)
DT 15-JUN-2002 (Ref. 41, Last annotation update)
DE Methylenetetrahydrofolate reductase 1 (Ec 1.5.1.20).
GN MET12 OR YPL023C OR LIP8C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansgore W.,
RA Araujo R., Aparicio A., Bartell B.G., Badcock K., Benes V.,
RA Bostein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duestelhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,

```

```

RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA Hummelke-Smith S., Hyman R., Johnston M., Kaiman S., Kleine K.,
RA Komp C., Kundl O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marache R., Messenguy F., Mewes H.-W., Mirzapat S., Moestl D.,
RA Mueller-Auer S., Namath A., Neutwich U., Oetner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schater M., Scharle M.,
RA Scherrens B., Schramm S., Schroeder M., Slicu A.M., Tetteijn H.,
RA Ureterazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Welsh S.V., Wambut R., Wang Y., Wedler E., Wedler H., Minnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.,
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI."
RL Nature 387:103-105(1997).
RN [2]
RP SEQUENCE OF 1-131 FROM N.A.
RX MEDLINE=85061207; PubMed=6095044;
RA Yang E., Friedberg E.C.;
RT "Molecular cloning and nucleotide sequence analysis of the
RT Saccharomyces cerevisiae Rad1 gene."
RL Mol. Cell. Biol. 4:2161-2169(1984).
RN [3]
RP IDENTIFICATION
RX MEDLINE=95004587; PubMed=7920641;
RA Goyette P., Sumner J.S., Milos R., Duncan A.M., Rosenblatt D.S.,
RA Mathews R.G., Rozen R.;
RT "Human methylenetetrahydrofolate reductase: isolation of cDNA,
RT mapping and mutation identification."
RL Nat. Genet. 7:195-200(1994).
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-
CC methylenetetrahydrofolate + NADPH.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: Folate metabolism.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; U36624; AAB68164.1; -
DR EMBL; K02070; -; NOT_ANNOTATED_CDS.
DR HSSP; P00394; 1B5T.
DR SCD; S0005944; MET12.
DR InterPro; IPR004621; fadh2_euk.
DR InterPro; IPR003171; Methydrif_redctse.
DR Pfam; PF02219; MTHFR; 1.
DR TIGRFAMs; TIGR00677; fadh2_euk; 1.
KW Oxidoreductase; Flavoprotein; FAD; NADP.
FT CONFLICT 110 111 DP -> NL (IN REF. 2).
FT CONFLICT 115 116 ED -> VV (IN REF. 2).
FT CONFLICT 119 119 ED -> VV (IN REF. 2).
FT CONFLICT 124 131 ESPFKYAV -> RLNLMLRF (IN REF. 2).
SO SEQUENCE 657 AA; 73941 MW; D75612BA92D67500 CRC64;

Query Match
Best Local Similarity 26.5%; Score 578.5; DB 1; Length 657;
Matches 151; Conservative 80; Mismatches 155; Indels 93; Gaps 16;

QY 1 ARVIVTQLFVDTDFLFKFNVDRCRQ-IGTCPIYVGPIMINNYKGFIRMTCCKTKIPADI 59
D 183 ADFITQIFVEPEAFIKFENFVRNHSNALRNIPPIPAIMPISQVSLKMTKRLCGCSVP 242
QY 60 MAALPE-IKDNEAVAKYGIHGTCEMKKI--LAHGITLHLYTNMKSALATIM--- 112
D 243 LSRFPPIQGDNAVAKSIGVDILIELIQETYSRIKGFHYTLNLKKAIAQIVSQSP 302
QY 113 -----NLGLI-----EESKVSRLPMRRPRAV 134
D 303 VLSHIYNESEEGEDETSGEIGSIENVPIEDADGIVLSDNSNETVANRRK--RRHSSL 360

```



DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).  
GN METF OR H11444.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Rd / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goeyne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Usterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT Rd.";  
RL Science 269:496-512(1995).  
CC -1- CATALYTIC ACTIVITY: 5-methylenetetrahydrofolate + acceptor = 5,10-  
CC methylenetetrahydrofolate + reduced acceptor.  
CC -1- COFACTOR: FAD (BY SIMILARITY).  
CC -1- PATHWAY: Methionine biosynthesis.  
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE  
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U32823; AAC23094.1; -.  
DR HSSP: P00394; 1B5T.  
DR TIGR: H11444; -.  
DR InterPro: IPR004620; Fadh2\_bact.  
DR InterPro: IPR003171; Methylot\_redctse.  
DR Pfam: PF02219; MTHFR; 1.  
DR TIGRFAMs: TIGR00676; fadh2; 1.  
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis;  
KW Complete proteome.  
SQ SEQUENCE 292 AA; 33020 MW; C6DCC9BC80CF9B4 CRC64;  
Query Match 8.6%; Score 189; DB 1; Length 292;  
Best Local Similarity 32.8%; Pred. No. 3.6e-08;  
Matches 38; Conservative 23; Mismatches 55; Indels 0; Gaps 0;  
Oy 1 ARVITOLFVYDFDIFLKVNDQIGITCPIVPGIMPINNYKGFIRMGFKTKIPADIM 60  
Db 175 ANHVITQFFPIENTLRKRDCAISGIDTEVPIGLPVTNKQIQKMAFNNVPIPMVLY 234  
Oy 61 AALEPIKNEBAVYAGIHGTETCKKILAHGIKTLHLVTLNNEKSALATIMNGL 116  
Db 235 KAYDGLDNDPTTNLVASVAMDMVKILSRGVNDFHFTILNRSELYALCHMLGV 290  
RESULT 12  
METF\_ECOLI STANDARD; PRT; 296 AA.  
AC P00394;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).  
GN METF OR B3941.

OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=84041480; PubMed=6356036;  
RA Saint-Girons I., Duchange N., Zakin M.M., Park I., Margarita D.,  
RA Ferrara P., Cohen G.N.;  
RT "Nucleotide sequence of metF, the E. coli structural gene for 5-10  
RT methylene tetrahydrofolate reductase and of its control region.";  
RL Nucleic Acids Res. 11:6723-6732(1983).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / M61655;  
RX MEDLINE=93347969; PubMed=8346018;  
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;  
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the  
RT region from 87.2 to 89.2 minutes."  
RL Nucleic Acids Res. 21:3391-3398(1993).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=99215588; PubMed=10201405;  
RA Guenther B.D., Sheppard C.A., Tran P., Rozen R., Matthews R.G.,  
RA Ludwig M.L.;  
RT "The structure and properties of methylenetetrahydrofolate reductase  
RT from Escherichia coli suggest how folate ameliorates human  
RL hyperhomocysteinemia.";  
RL Nat. Struct. Biol. 6:359-365(1999).  
CC -1- CATALYTIC ACTIVITY: 5-methylenetetrahydrofolate + acceptor = 5,10-  
CC methylenetetrahydrofolate + reduced acceptor.  
CC -1- COFACTOR: FAD.  
CC -1- PATHWAY: Methionine biosynthesis.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE  
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: V01502; CAA24747.1; -.  
DR EMBL: U19201; AAB03073.1; -.  
DR EMBL: AE000468; AAC76923.1; -.  
DR PIR: A00462; RDECMH.  
DR PIR: S40884; S40884.  
DR PDB: 1B5T; 20-JAN-99.  
DR BioGene: BG10585; metF.  
DR InterPro: IPR004620; Fadh2\_bact.  
DR InterPro: IPR003171; Methylot\_redctse.  
DR Pfam: PF02219; MTHFR; 1.  
DR TIGRFAMs: TIGR00676; fadh2; 1.  
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis;  
KW 3D-structure; Complete proteome.  
SQ SEQUENCE 296 AA; 33102 MW; B702702D2BE9521E CRC64;  
Query Match 8.6%; Score 189; DB 1; Length 296;  
Best Local Similarity 32.1%; Pred. No. 3.7e-08;  
Matches 36; Conservative 25; Mismatches 51; Indels 0; Gaps 0;  
Oy 5 VVGLFVDFDIFLKVNDQIGITCPIVPGIMPINNYKGFIRMGFKTKIPADIMALE 64  
Db 181 ITQFFDVESYILNFRDRCVAGIDVEIIPGILPVSNFQAKKFAADWTNVRIPAMMAQWD 240  
Oy 65 PINDNEBAVYAGIHGTETCKKILAHGIKTLHLVTLNNEKSALATIMNGL 116  
Db 241 GLDDDAETKLVGANIAMDMVKILSRGVNDFHFTILNRSELYALCHMLGV 292

```

RESULT 13
METF_SALTY STANDARD; PRT; 296 AA.
ID METF_SALTY STANDARD; PRT; 296 AA.
AC P11003;
DT 01-JUL-1989 (Rel. 11, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
GN METF OR STM4105.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=88302115; PubMed=2841568;
RA Stauffer G.V., Stauffer L.T.;
RT "Cloning and nucleotide sequence of the Salmonella typhimurium LT2
RT metf gene and its homology with the corresponding sequence of
RT Escherichia coli.";
RL Mol. Genet. 212:246-251(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SSGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + acceptor = 5,10-
CC methylenetetrahydrofolate + reduced acceptor.
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: Methionine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X07689; CAA30531.1; -.
DR EMBL: AE008891; AML22945.1; -.
DR PIR: S03169; S03169.
DR HSSP: P00394; 1B5T.
DR StyGene: SG10224; metf.
DR InterPro: IPR004620; Fadh2_bact.
DR InterPro: IPR003171; Methylcrof_redctse.
DR Pfam: PF02219; MTHFR; 1.
DR TIGRFAMs: TIGR00676; fadh2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis;
KW complete proteome.
FT CONFLICT 94 94 A -> P (IN REF. 1).
SQ SEQUENCE 296 AA; 33173 MW; 2B834F80B56A643 CRC64;

Query Match 8.4%; Score 184; DB 1; Length 296;
Best Local Similarity 31.2%; Pred. No. 9.4e-08;
Matches 35; Conservative 26; Mismatches 51; Indels 0; Gaps 0;

OY 5 VTQLFYDTDLKLVNDRCQIGITCPVIGIMPINNYKGFIRMTGFCRTKIPADIMAALE 64
DB 181 ITQFFDVESYLRFRDRCVAGSAGDVEIIGILPVSNFKQAKKADMTNVRIPSSWMSIMRE 240
OY 65 PIKDNEAVKAYGIHLGTCKKILAHGKIKTLHLTYLNMKSALAILMNLGL 116

```

```

DB 241 GLDNDAEFRKLVGANIAMDWKILSRGVKDFHRYTLNRAEMSYAICHTLGV 292
: : : | | | : : | : | | | | : : | |
RESULT 14
METF_ERWCA STANDARD; PRT; 298 AA.
ID METF_ERWCA STANDARD; PRT; 298 AA.
AC P71319;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
GN METF.
OS Erwinia carotovora.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pectobacterium.
OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=71;
RA Calcutt M.J., Lewis M.S., Eisenstark A.;
RX Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + acceptor = 5,10-
CC methylenetetrahydrofolate + reduced acceptor.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: Methionine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U74302; AAC72242.1; -.
DR HSSP: P00394; 1B5T.
DR InterPro: IPR004620; Fadh2_bact.
DR InterPro: IPR003171; Methylcrof_redctse.
DR Pfam: PF02219; MTHFR; 1.
DR TIGRFAMs: TIGR00676; fadh2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis.
SQ SEQUENCE 298 AA; 33589 MW; 50FE729E4E3E2C56 CRC64;

Query Match 8.0%; Score 176; DB 1; Length 298;
Best Local Similarity 30.4%; Pred. No. 4.3e-07;
Matches 34; Conservative 25; Mismatches 53; Indels 0; Gaps 0;

OY 5 VTQLFYDTDLKLVNDRCQIGITCPVIGIMPINNYKGFIRMTGFCRTKIPADIMAALE 64
DB 181 ITQFFDVESYLRFRDRCVAGSAGDVEIIGILPVSNFKQAKKADMTNVRIPSSWMSIMRE 240
OY 65 PIKDNEAVKAYGIHLGTCKKILAHGKIKTLHLTYLNMKSALAILMNLGL 116
DB 241 GLDNDPETRKRVKVASIAMDVKIILSRGVKDFHRYTLNRAELSYAICHTLGV 292
: : : | | | : : | : | | | | : : | |
RESULT 15
METF_BUCAL STANDARD; PRT; 292 AA.
ID METF_BUCAL STANDARD; PRT; 292 AA.
AC P57154;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
GN METF OR BU046.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OC symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-TOKYO 1998;
RX MEDLINE-20445173; PubMed-10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RL Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -! CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + acceptor = 5,10-
CC methylenetetrahydrofolate + reduced acceptor.
CC -! COFACTOR: FAD (BY SIMILARITY).
CC -! PATHWAY: Methionine biosynthesis.
CC -! SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001118; BAB12769.1; .
DR HSSP; P00394; 1B5T.
DR InterPro; IPR004620; Fadh2_bact.
DR InterPro; IPR003171; Methylcrof_redctse.
DR Pfam; PF02219; MTHFR; 1.
DR TIGRFAMs; TIGR00676; fadh2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis;
KW Complete proteome.
SQ SEQUENCE 292 AA; 33730 MW; 1D2A9E1B4D6E0DA9 CRC64;

Query Match 7.7%; Score 168; DB 1; Length 292;
Best Local Similarity 30.4%; Pred. No. 1.9e-06;
Matches 34; Conservative 24; Mismatches 54; Indels 0; Gaps 0;

OY 5 VTQLFYDTDFLKLKVNDCRQIGITCPIVPGIMPINNYKGIPTMTGFCCKTIPADIMALE 64
Db :|||: :||: :||: | :|||:|: ||: ||: :
181 ITOEFFNIESYLRFDNCIKRKNINDITPGILPVCNFOKLRFSMTNVKIPKMLDMFN 240
OY 65 PIKDNIEAVKAYGTHLGTEMKCKILAHGIKTLHLTYLTMKKSALAILMNLGL 116
Db :||: | :||: :||: ||| :|| :||
241 GLDDDIPTQKITIGSSIALDMVKKLSCEGVKNFHEHTLNOSDITYSICHIIGL 292

```

Search completed: February 14, 2003, 23:08:18  
 Job time : 25 secs



# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 2/24/03  
Searcher: P. Schreiber  
Terminal time: 18  
Elapsed time: 11  
CPU time: \_\_\_\_\_  
Total time: \_\_\_\_\_  
Number of Searches: \_\_\_\_\_  
Number of Databases: \_\_\_\_\_

### Search Site

\_\_\_\_ STIC  
☒ CM-1 6A03  
\_\_\_\_ Pre-S

### Type of Search

2 N.A. Sequence  
1 A.A. Sequence  
\_\_\_\_ Structure  
\_\_\_\_ Bibliographic

### Vendors

\_\_\_\_ IG  
\_\_\_\_ STN  
\_\_\_\_ Dialog  
\_\_\_\_ APS  
\_\_\_\_ Geninfo  
\_\_\_\_ SDC  
\_\_\_\_ DARC/Questel  
☒ Other Compass

